Performance of Adaptive Admission/Congestion Control Policies on a Hybrid MC-CDMA/TDMA Integrated Multimedia Networks

Uthman A. Baroudi

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy at

Concordia University

Montréal, Québec, Canada

September 2000

CONCORDIA UNIVERSITY

Division of Graduate Studies

This is to certify that the thesis prepared

By: **Uthman A. Baroudi**

Entitled: **Performance of Adaptive Admission/Congestion Control Policies on a Hybrid MC-CDMA/TDMA Integrated Multimedia Networks**

and submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

______________________ Dr

______________________ Examiner A

______________________ Examiner B

______________________ Examiner D

______________________ Dr. Ahmed K. Elhakeem

Approved by ______________________

________ 20 ____ ______________________

Dean of Faculty

ABSTRACT

Performance of Adaptive Admission/Congestion Control Policies on a Hybrid MC-CDMA/TDMA Integrated Multimedia Networks

Uthman A. Baroudi, Ph.D.
Concordia University, 2000

Future 3G CDMA networks are expected to have the ability to accommodate variety of users, each with its own transmission characteristics and quality of service (QoS) requirements to be maintained. Compatible multi-access system should provide the means to control (i.e. admission/congestion) the flow of traffic and at the same time maintain the QoS requirements

The essence of this work is to introduce an interaction between the physical layer and higher network layers, thus enabling more practical utilization of multi-user detection and supporting services with different QoS parameters. The traffic sources are classified according to their activities and QoS of parameters into two categories, namely stream and interactive traffic. As congestion arises, the interactive traffic users are granted higher priority than the stream traffic users. In this work, a new hybrid Multicode (MC)-CDMA/TDMA medium access control technique utilizing multi-user detection accompanied with a traffic flow control scheme is proposed and analyzed.

Secondly, the End-to-End performance criteria of the proposed MAC such as packet delay, channel utilization, etc. are investigated under new Window Measurement-Based admission/congestion control policy where the status of the buffer at the base

station is estimated taking into consideration the physical limitations of the base station (i.e. the number of transmission and reception modems), call and burst level traffic, channel impairment, etc.

Thirdly, the hybrid analysis/simulation routine is repeated to investigate and verify the performance of the above proposed MAC along with several versions of congestion policies. The whole system is studied under wide range of traffic load, and different system parameters. The results show that the QoS requirements have been effectively maintained and the proposed system is amenable to practical implementation.

SOMMAIRE

Performance of Adaptive Admission/Congestion Control Policies on a Hybrid MC-CDMA/TDMA Integrated Multimedia Networks

Uthman A. Baroudi

To My Beloved Paraents

ACKNOWLEDGEMENTS

All praises be to ALLAH (GOD), The Sustainer, The Most Merciful The Most Beneficent Who I feel His support and Help in every instance of my life. Then, I would like to express my deep thanks and gratefulness to my thesis supervisor Prof. Elhakeem for the patient support throughout all phases of this work. His guidance and instructions are always source of inspiration and motivation for the completion of this research in its current shape.

I am also very thankful for the examination committee for their fruitful suggestions and giving their precious time to read and examine my work.

I am indebted to my colleagues who have been very supportive and giving me the courage to complete this work.

Finally, special thanks to my parents for their constant support and prayers. I am in a real debt to my wife who was beside me throughout all my Ph.D. study. Her great effort in taking care of our children plays a key role in preparing this work. I seek the forgiveness of my family especially my children for not giving them much of my time, but I promise them I will give them the best in the future.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

$\hat{b}$	Estimated transmitted bit
$O(\cdot)$	Order of computation complexity
$\mathcal{R}$	kxk Cross-correlation matrix
τ_k	Transmission delay
T_b	Bit duration
N_o	White noise single-sided power spectral density
$\rho_{i,k}$	Cross-correlation between ith and kth signals
$P_e(k)$	Probability of error
P	Transmitted signal power
W	Bandwidth
η	Background noise power
I_o	Total noise power
σ	standard deviation
SIR	Signal-to-Interference Ratio
SIR_o	Minimum SIR for proper operation
SIR_{TH}	Threshold SIR
$\hat{\xi}$	Ratio of a user's average bit-rate to the basic rate
ξ	Ratio of a user's peak bit-rate to the basic rate
$\theta_{j,on}$	The packet level activity of a user from class j
β	Rate transition out of silent state
α	Rate transition out of talk spurt
λ_i	Compound probability distribution of i packets
i_p^k	Instantaneous number of packets belonging to all high priority traffic classes
p_{ij}	Transition probability from state i to state j

γ_j	Probability of generation j packets and successfully received via uplink channel
u_j	Probability of successfully reception of j packets
μ	Average service rate, packet/sec
$O_b(\cdot)$	Buffer overflow probability
L	Maximum number of servers on the downlink channel
B	Base station buffer size

Chapter 1

Introduction

1.1 Motivation

The telecommunication community is moving quickly to build an infrastructure for the Personal Communication Networks (PCNs). This technology aims to integrate voice, video, image, and data in both telecommunication and computing environments. Hence, PCNs should be capable to provide a wide range of varieties of services such as high speed transmission, flexible bandwidth allocation, variable bit rate (VBR), constant bit rate (CBR), available bit rate(ABR), quality of service (QoS) selections [1].

The question is how effectively should the resources at our disposal be utilized such that the above objective is accomplished? The answer for this question is not an easy task. Of course, the focus, here, is on the frequency, time and space resources. Radio spectrum has the dimensions of frequency, time, and space. The same frequency may be used for different purposes at the same time in sufficiently separated spaces (e.g. cellular system), the same frequency may be used for different purposes in the same space at sufficiently separated times, and sufficiently separated frequencies may be used at the same time and at the same place. Thus the reduction of bandwidth for transmission means more frequency bands would be available for other purposes at the same place; the reduction of physical space for transmission would allow the same frequency band to be used for different purposes at the same time in more spaces; the reduction of transmission time would allow the frequency band to be used for more purposes in the same space. As a result, the reduction of spectrum space used to transmit a given quantity of information can increase the spectrum efficiency [7]. The interesting problem arising here is the interference from other users who are using the same common channel. Hence, to achieve successful transmission, interference must be avoided or at least controlled [9]. The schemes that perform such duties are called Multiple Access Protocols. From the above discussion, we can view the following basic multiple access protocols: Frequency Division Multiple Access (FDMA), Time Division Multiple Access (TDMA), and

Code Division Multiple Access (CDMA).

The CDMA system is considered as one of the very promising candidates for wireless ATM [1] because of its many attractive features. To mention few, first, the CDMA networks are well known to combat effectively frequency-selective fading, jamming, and other in-band interference [2]-[4]. Second, they easily enable the integration of heterogeneous traffic users [5].

However, there are three main disadvantages in the CDMA protocol that might compromise the above mentioned advantages. First, CDMA protocol does not easily support high bit rate users which is very very important for future traffic. For example, let the available bandwidth be 20 MHz (which is the maximum proposed bandwidth for future **Wideband CDMA** systems [6]), and the processing gain 100, then the maximum bit rate that can be supported is 0.2 Mb/s. Second, the Direct-Sequence CDMA (DS-CDMA) scheme requires tight power control to achieve good performance.

Third, it is well known that CDMA communication systems are interference limited. These interferences are due to other users who are sharing the same frequency channel simultaneously. Hence, the detection process in such environment is not straight forward, because the desired signal is corrupted by other signals which are intended by other users. Therefore, the interference consists of two main parts: the interferences due to other users and the additive white Gaussian noise. This interesting situation motivates the researchers to propose several scenarios for the CDMA receiver. The proposed receivers can be classified into two classes: *Single-user* receiver (Conventional) and *Multi-user* receiver. This classification refers to the way the receiver should treat or deal with undesired signals. The later jointly detects the desired signals using the fact that the receiver already has information about the transmitted signals such as the signature code of each user or it can estimate important characteristics of the transmitted signals (such as energy of the received signals) that can help in improving the process of jointly detecting these signals. On the other hand, the conventional single user receiver treats the received signal as if it is composed of the desired signal plus noise (e.g. Gaussian Noise). Therefore, this receiver deals with each user separately.

Further, the seminal work on multiuser detection by Verdu [41] has added another dimension to the expected future communication systems. The theoretical performance of multiuser detection shows much improvement compared to the conventional systems (single-user). Yet, the optimum multiuser detector has a prohibitive complexity which makes it undesirable for wireless communication. This motivates researchers to look for sub-optimum multiuser detectors (see e.g. [62, 42]) that have good performance and at the same time affordable complexity. Nevertheless, in multiuser detection, it is generally assumed that the number of simultaneous packets, the identities of the users' codes, and the multipath channel associated with each user transmission are perfectly known. Deviating from such assumptions will yield a very high probability of error. For connection-oriented services, the signaling period may yield useful information about the number of users. Unfortunately, even with such knowledge, users burstiness will preclude the accurate estimation of the instantaneous number of packets on the channel.

Now, it should be clear that designing a multiple access control depending mainly on physical layer parameters without interaction with other layers in the networks would lead into improper and inefficient usage of the resources(e.g. bandwidth, servers, time, etc.). Therefore, introducing an interaction between the access protocol and the activity of the users as well as its requirements of service quality is very essential for reliable and practical networks. Most of previous works (e.g. [72], [77], [87]) focused on the "worst case" scenario assuming fixed traffic population. For instance, in [77], system capacity is evaluated under the assumption of fixed number of calls, fixed QoS parameters and so on. As a matter of fact, the actual activity of the network is far from such rigid parameters. Nowadays and in the near future, the nature of the traffic population for a wireless network is much diversified.

Examining the *Decorrelator Receiver* [8], we can see many attractive features that make it viable solution to future CDMA networks. First of all, the multiuser interference has been eliminated. Second, its computational complexity is much reduced compared with the optimum receiver. The computational complexity of the decorrelator detector increases linearly with the number of users (i.e. $O(K)$). Third, it uses a linear transformation to obtain an estimate of the transmitted symbol. At practical SNR values, the decorrelator receiver provides much better performance than the conventional one. Further, the linear decorrelator receiver exhibits the same degree of near-far resistance as the optimum multiuser detector. In addition, when the users energies are unknown, the decorrelator receiver is the optimal approach [62]. However, a significant limitation of this technique is the computational complexity due to the inversion of the correlation matrix [45] where entries depend on the number of active users, signature sequences, and the delays of the users. Many researches have tried minimizing the computation required, for example Ref. [45]. Further, any change in one of these parameters changes the correlation matrix and consequently a need for updating the multiuser detection process. Moreover, the uncertainty in the actual number of active users is another serious problem that might degrade the system performance very severely [66] which now could be resolved by applying our proposed scheme for traffic control.

Motivated by the discussion above, the essence of this work is to introduce an interaction between the physical layer and higher layers, thus enabling a more practical utilization of multiuser detection and supporting services with different QoS parameters. To achieve this objective, the traffic sources are classified according to their activities and QoS of parameters into two categories, namely stream and interactive traffic. In the case of contention, the interactive traffic users are granted higher priority than the stream traffic users. In this work, a new hybrid Multicode (MC)-CDMA/TDMA medium access control utilizing multiuser detection is proposed and analyzed. Further, two traffic flow control approaches, to accompany the

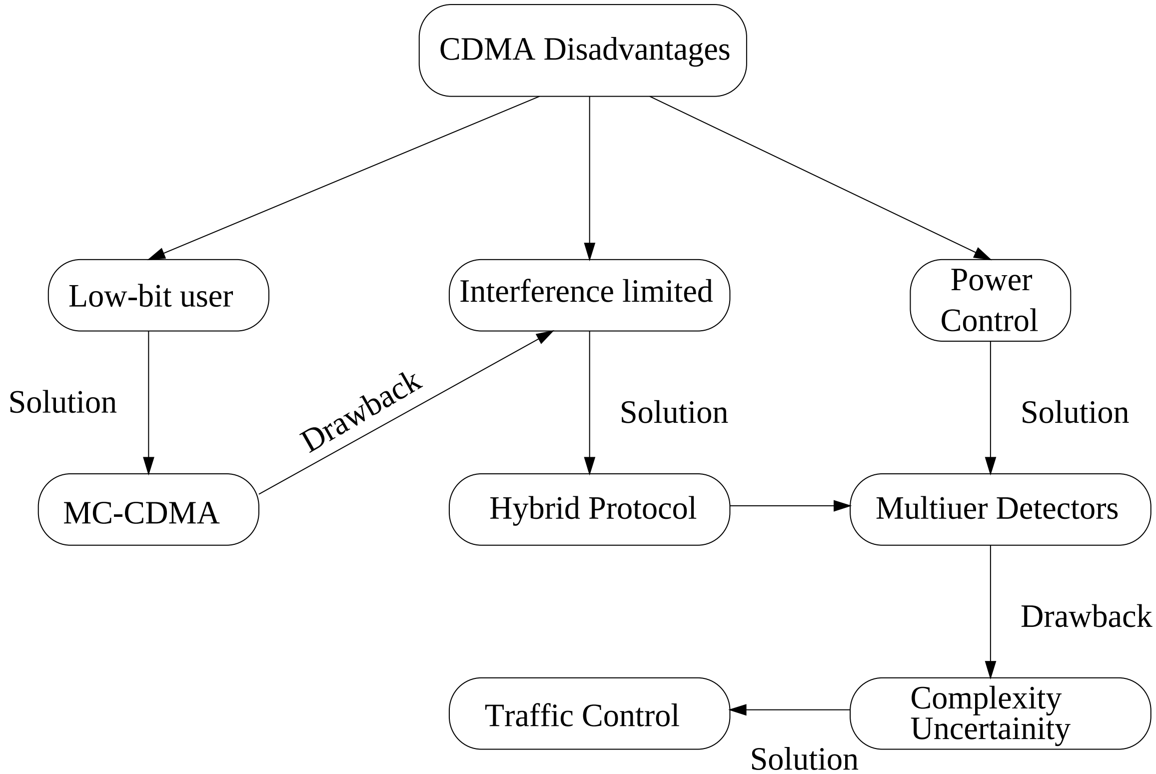


Figure 1.1: Motivation behind the proposed access protocol

TDMA/MC-CDMA system, are proposed. One approach deterministically controls the flow of traffic into the TDMA slots, while the other statistically controls the flow of traffic depending on the instantaneous changes in the traffic load. Although, hybrid TDMA/CDMA techniques were proposed in the literature before (see e.g. [12]), to our knowledge, our proposal is new and original at least by creating an interaction between the physical layer and other higher layers (for example, ATM layer and or the flow and congestion control functions in the data links or network layers) that could enable the integration of multimedia applications into wireless networks.

Figure 1.1 illustrates the motivation behind our proposed hybrid MC-CDMA/TDMA access protocol. It reads that CDMA systems can support high bit rate users if *Multicode* MC-CDMA techniques used and consequently we solve a problem and at the same time creating another serious problem that is high flux of mutual interference of low-bit rate streams. The proposed solution is to adopt a hybrid access protocol that could balance the traffic flow. More, looking at the tight power control issue, we could use *Decorrelator* receiver to lessen the impact of this drawback of DS-CDMA protocol. Nevertheless, as mentioned above, *Decorrelator* receiver has computational complexity and its theoretical performance depends highly on the knowledge of the exact number of active users communicating with the base station. Our proposed remedy is to impose a control scheme on the flow of traffic into the network such that the base station has 'perfect' knowledge about the number of packets processed by the *Decorrelator* multiuser receiver.

Applying the above proposed access protocol shows outstanding performance in terms of **BER** and in balancing heterogeneous traffic between TDMA slots. Nevertheless, there is another important issue that should be addressed that is the *Admission / Congestion* control issue. In other words, we should provide answers or clues on the following questions. First, does the new call affect the QoS of active callers currently carried by the network? Second, can the network provide the QoS required by the new caller? The question then becomes how can we map the end-to-end QoS requirements into physical parameters. In other words, how can we map the link layer QoS into physical layer QoS [76]. What is the mechanism that should be followed? Is it in implementing *preventive* congestion control where the schemes are designed to prevent the occurrence of congestion? or should we apply *reactive* congestion control where one relies on the feedback information for controlling the level of congestion? [67]

Many researches have been published addressing this issue (e.g. see [72]-[77]). Yet, most of the existing work considers either the uplink or the downlink channel in evaluating the CDMA system under the proposed admission/congestion policies. Further, These studies do not include the error performance in the queueing modeling of the CDMA networks. Nevertheless, the inclusion of the channel error performance in the queueing modeling would give more realistic figures of performance as well as it guarantees the QoS of the networks. Therefore, the second part of this work shall focus on devising an adaptive call admission/congestion policy based on a Window-Measurement estimation of the status of the buffer at the base station. In our analysis, we interrelate physical limitations of the base stations (i.e., the number of transmission and reception modems), call and burst level traffic, instantaneous BS buffer condition, and End-to End bit error performance in one queueing problem. Then, a Window-Measurement estimator is developed to estimate the probability that the buffer at the base station could be congested and accordingly, the traffic load shall be controlled.

1.2 Scope of the Thesis

This thesis investigates a proposed interaction between the design of the physical layer and the nature of the application that are using the networks. The related technologies are too wide to be covered. Although, we are going to briefly cover the state-of-the-art related to three aspects that are *Multiple access protocols, Multiuser receivers and admission/congestion policies for CDMA networks.*

We begin with Chapter 2 where many multiple access protocols are introduced including the basic ones such as FDMA, TDMA and CDMA. First, we shall discuss briefly each of the above basic multiple access techniques. Then, we shall explore the influence of the traffic changes on the choice of multiple access. Finally, the hybrid multiple access techniques are addressed. The implication of the multiple access techniques on the wireless networks are discussed along with our introduction to each protocol.

Chapter 3 covers the main classes of receivers proposed in the literature for direct sequence CDMA (DS-CDMA) with more emphasis on *Decorrelator Multiuser* receiver which is a key element of of our proposal.

Finally, we end these brief introductory chapters with chapter 4 which explores the existing work in the area of Admission/Congestion control policies for CDMA networks.

Our contribution starts at chapter 5 where a new hybrid TDMA/ Multicode (MC)-CDMA medium access control utilizing multiuser detection is proposed and analyzed. Further, two traffic flow control approaches to accompany the TDMA/MC-CDMA system are proposed. One approach deterministically controls the flow of traffic into the TDMA slots, while the other statistically controls the flow of traffic depending on the instantaneous changes in the traffic load. The two approaches have been examined under a wide range of traffic characteristics where AWGN is only considered besides the mutual interferers from other intracell users. Both approaches show superiority as well as less sensitivity in terms of BER to the traffic changes compared with the conventional system.

Chapters 6 and 7 contain the second part of this work which focuses on devising an adaptive call admission/congestion policy based on a Window-Measurement estimation of the status of the buffer at the base station.

The complexity of the wireless networks under consideration has put some restrictions on our analysis. Hence, it is important to evaluate the performance of the proposed protocols via simulation where these protocols shall be tested under more realistic conditions and system parameters. Chapter 8 covers the simulation set-up for exact Monte-Carlo simulation used to simulate the multimedia integrated CDMA networks where heterogeneous traffic users are multiplexed into simple TDMA frames. In this simulation, our emphasis is on the network performance issues such as packet delay, packet delay jitter, frame losses, call blocking, etc. Therefore, the only system parameters that are going to be simulated and randomly generated throughout the simulation program are these parameters bearing direct relation to the network performance. Then, the simulation results are presented and discussed. We end this work by chapter 9 with main conclusions and suggestions for future work.

Chapter 2

Multiple Access Techniques

2.1 Introduction

How effectively the resources at our disposal should be utilized? This question faces us in every aspect in our life. The answer for this question is not an easy task. Of course, the focus, here, is on the frequency, time and space resources. Radio spectrum has the dimensions of frequency, time, and space. The same frequency may be used for different purposes at the same time in sufficiently separated spaces (e.g. cellular system), the same frequency may be used for different purposes in the same space at sufficiently separated times, and sufficiently separated frequencies may be used at the same time and at the same place. Thus the reduction of bandwidth for transmission means more frequency bands would be available for other purposes at the same place; the reduction of physical space for transmission would allow the same frequency band to be used for different purposes at the same time in more spaces; the reduction of transmission time would allow the frequency band to be used for more purposes in the same space. As a result, the reduction of spectrum space used to transmit a given quantity of information can increase the spectrum efficiency [7].

Following the above philosophy, the literature has a considerable amount of available research trying to find ways complying with this philosophy. In this proposal, we will explore application of the above philosophy on the multiple users and multiple communication links rather than point-to-point links.

There are several types of multiuser communication systems, such as, a multiple access system, a broadcast network, a store-and-forward network. In multiple access systems, a common channel is accessed by a large number of users to transmit information to a receiver (up-link in satellite system or the up-link in the mobile cellular network) as shown in fig. 1. In the broadcast system, we have the opposite where a single transmitter sends information to multiple receivers (down-link in the mobile cellular network) [8].

The interesting problem arising here is the interference from other users who

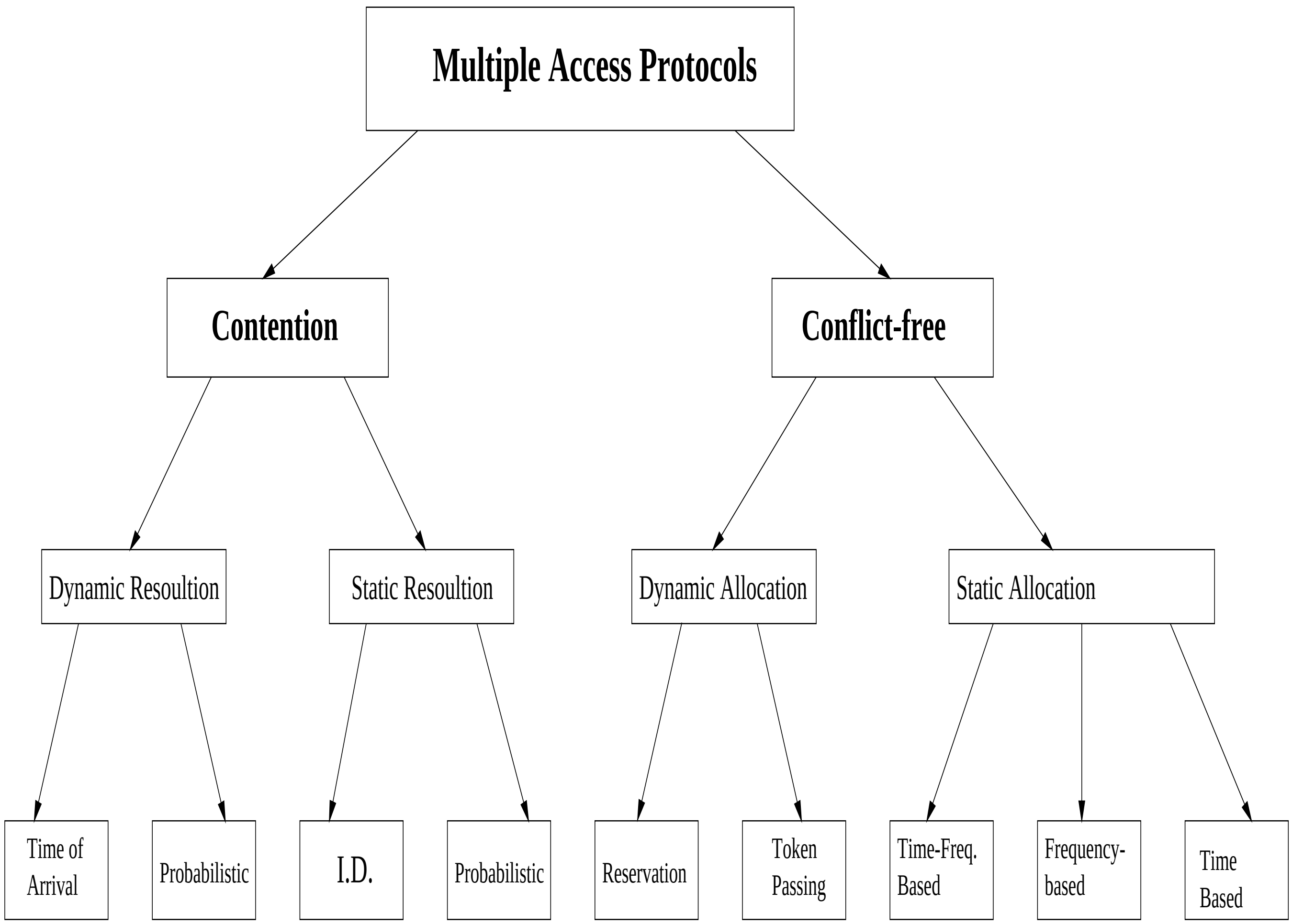


Figure 2.1: Classification of Multiple Access Protocols [9]

are using the same common channel. Hence, to achieve successful transmission, interference must be avoided or at least controlled [9]. The schemes that perform such duties are called Multiple Access Protocols. These protocols can be classified into two basic classes, namely, conflict-free and contention (Random Access) protocols. Conflict-free protocols ensure that a transmission, whenever made, is a successful one, that is, will not be interfered by another transmission. This objective can be achieved by allocating the common channel to the users either statically or dynamically [9]. The main drawback of this class is that, under heavy population traffic, this protocol is not any more effective. The alternative to the conflict-free protocol is the contention protocol where every user is allowed to transmit, but the transmission is not guaranteed to be successful, because of the expected collision among the simultaneous active users. Two packets are considered in collision, if they were transmitted during the same time slot and they could not be received successfully. Hence, the protocol must prescribe a way to resolve conflicts once they occur so all messages are eventually transmitted successfully. Again, the resolution is either static or dynamic [9]. Fig. 2.1 illustrates the general classification of multiple access protocols.

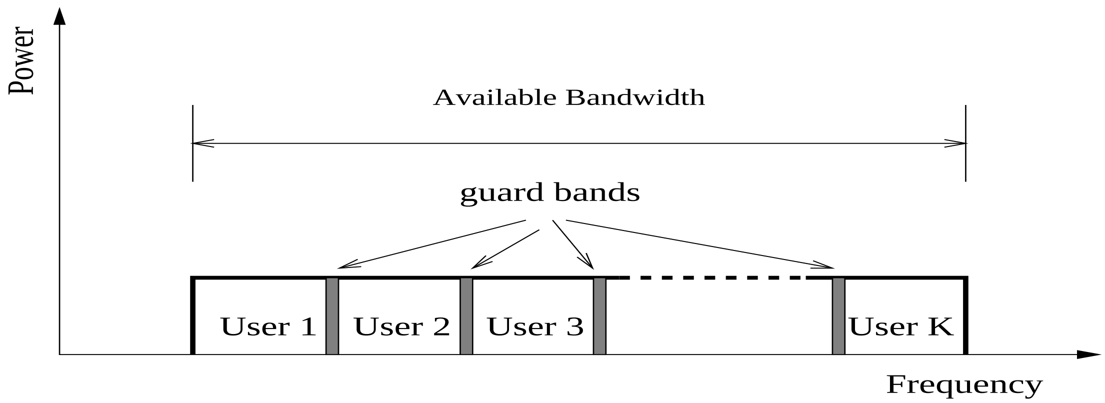
Another way to look at the problem (efficient spectrum utilization) is to consider the traffic population variations, and accordingly we classify the ways that may accommodate such traffic. Here, we shall follow this approach of classification. In other words, our discussion for the existing multiple access protocols will be according to their compatibility to the traffic requirements.

From the above discussion, we can view the following basic multiple access protocols: Frequency Division Multiple Access (FDMA), Time Division Multiple Access (TDMA), Code Division Multiple Access (CDMA). First, we shall discuss briefly each of the above basic multiple access techniques. Then, we shall explore the influence of the traffic changes on the choice of multiple access. Finally, the hybrid multiple access techniques are addressed. The implication of the multiple access techniques on the wireless networks are discussed along with our introduction to each protocol.

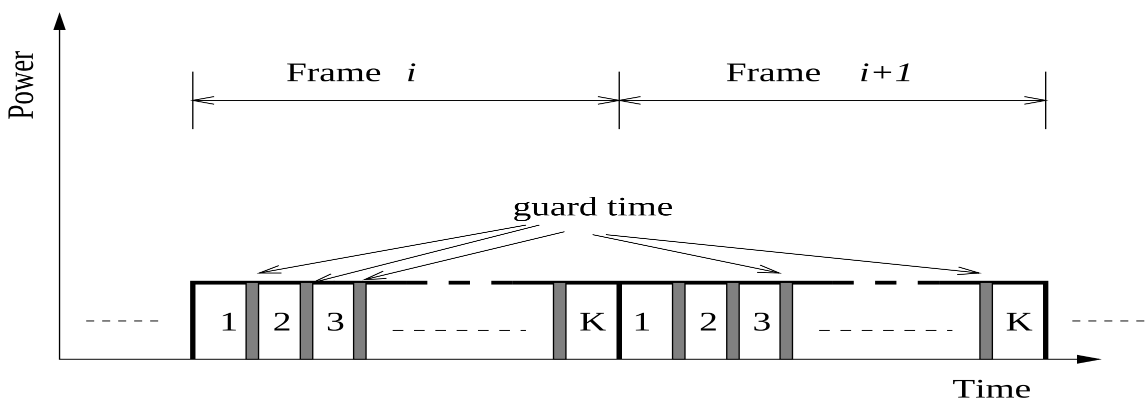
2.2 Basic Multiple Access Techniques

2.2.1 Frequency division Multiple Access (FDMA)

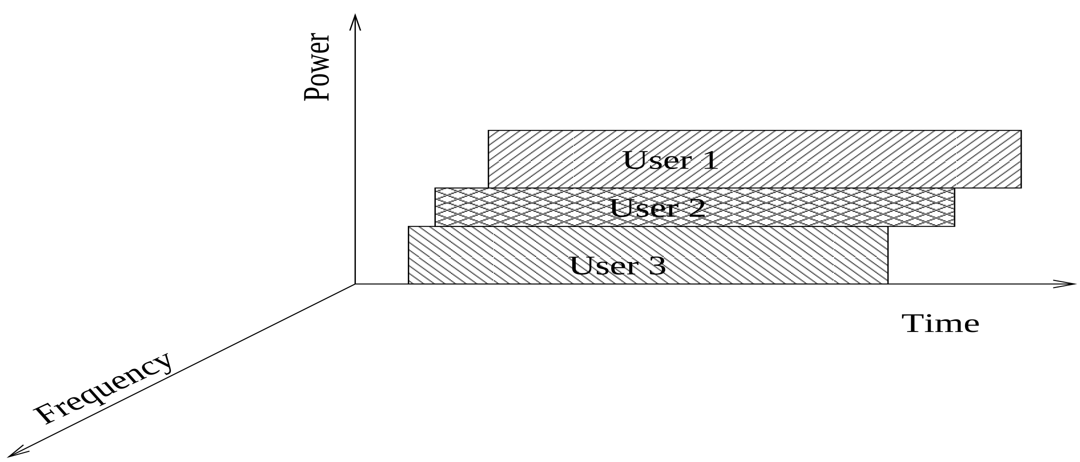
The basic principle of FDMA is that the entire available bandwidth is divided into bands each of which serves a single user as illustrated in Fig. 2.2a. It is the only multiple access technique that can be used in analog transmission as well as digital [99]. The attractive feature of this protocol is its simplicity. It does not require any time synchronization among the users, because each user has its own frequency spectrum to use without any interference (ideally) from other users [9],[16]. In practice, FDMA protocol suffers several problems, for instance, intermodulation effects, adjacent channel interference. Therefore, guard bands must be used between adjacent frequency channels to minimize these effects. Consequently, the frequency utilization efficiency is reduced. The required size of guard band depends in part on the residual sidebands in each transmitted signal [16]. This could explain why the primary task for radio engineers, for decades, was to design filters that can tackle the problem of separating the desired user's spectrum from other users' traces who occupy other frequency bands [99]. For more details about the above issues in FDMA and also the system analysis, the reader is directed to Ref. [16], [100]. . Further, in wireless networks or in personal communications where the traffic is expected to be high, the critical issue, here, is what is known as channel assignments. The channel assignment is a protocol that controls the way of assigning frequency band for each user. This can be done statically or dynamically. This issue shall be discussed later.



a) FDMA Protocol assignments for K users.



b) TDMA Protocol assignments for K users.



c) Asynchronous CDMA Protocol.

Figure 2.2: Basic Multiple access Techniques

2.2.2 Time Division Multiple Access (TDMA):

This protocol tries to efficiently use the time resource. In TDMA, the time resource is divided into time slots, where each user is granted one of these slots, while the whole bandwidth can be used by this user. In TDMA, the bit rate of the transmitted bursts are generally many times higher than that of the continuous input bit streams. The slot assignments follow a predetermined pattern that repeat itself periodically; each such period is called a cycle or a frame [9],[16]. Hence, a buffer is needed at each transmission to store the data bits received from one frame until the next. Fig. 2.2b illustrates the TDMA protocol. This multiple access protocol imposes the following constraints on the system design: 1) time synchronization, 2) buffer design. Further, guard times should be introduced between users' slots to avoid inter-modulation (IM) products and also to accommodate any timing inaccuracies. On the other hand, TDMA is the most helpful protocol in supporting high peak rate users [16],[30],[37].

2.2.3 Code Division Multiple Access (CDMA):

This technique has been known since about 1940 [15]. CDMA employs the spread spectrum modulation format, meaning that each user's digital waveform is spread overall the entire spectrum allocated to all users of the network [99]. The spreading function is done according to a code which is known by both the transmitter and the receiver. In principle, each user is assigned different code and because of that the technique is Code Division Multiple Access. In the receiver side, the intended user despreads the received signal using his information about the spreading code and then demodulate the user's digital signal. Of course, the other signals intended to other users remain spread with very low power spectral density. From the above discussion, it is obvious that CDMA has no time restriction as the case in TDMA nor frequency restriction as in FDMA. Fig. 2.2c illustrates the asynchronous CDMA.

Code division multiple access [39] can be classified into three classes according to the characteristics of the spread spectrum signal: 1) Direct Sequence CDMA (DS-CDMA) where the spectrum of the data-modulated signal is widened directly by a second modulation using a wide band signal or code, 2) Frequency Hopping CDMA (FH-CDMA) where the widening of the spectrum is accomplished by changing the carrier frequency periodically and each carrier frequency is chosen from a set of frequencies which are spaced approximately the width of the data modulation bandwidth apart, and 3) Time Hopping CDMA (TH-CDMA) is the counterpart to FH-CDMA where the time domain is divided into frames and each frame is further split into M time slots. During each frame one and only one time slot will be modulated with a message by any reasonable modulation method. The particular time slot chosen for a given message is selected randomly. For more details, the reader is referred to Ref [39],[40].

2.3 Traffic Dynamics and the Choice of Multiple Access Protocol

Having discussed the basic multiple access techniques, the question is how to choose the proper protocol? In general, there are two primary factors that direct the choice of a multiple access technique: 1) the traffic characteristics of the network of interest, and 2) the state of technology development at the time a network is deployed [35]. When we look at the first factor (traffic characteristics), we have to observe three things. First, the nature of traffic (i.e steady, dynamic). Second, the transmission characteristics of the traffic population (i.e. CBR, VBR, ABR, etc.). Third, The quality of service (QoS) requirements by the end user.

Considering a communication network, where the traffic from each user is steady, or nearly so, it is possible to assign a fixed portion of the multiple-access channel (Fixed channel assignment) for each user. The ordinarily chosen candidates for such protocol is either TDMA, or FDMA. As long as the traffic of each user is relatively stable, FDMA and TDMA provide acceptable performance [35]. However, in many networks, we do not have the above situation, but we observe that the traffic from each user varies with time. Furthermore, the number of active users also changes as function of time [10],[35]. This bursty nature of the traffic manifests itself either in the call level or in the packet level. Accordingly, the time slot or the frequency band assigned to a certain user is not used at the moment this user is idle, while there might be other users who are in need for this slot and they are denied to have an access to.

Therefore, it is clear that the fixed channel assignment technique considered above does not fully exploit the capabilities of the system, because of the poor utilization of the available bandwidth due to the variations in the traffic [10]. In this case, it may be desirable to consider the Demand Assigned Multiple Access (DAMA) architecture [35]. It dynamically assigns channel according to the demand of each user and the traffic load [10]. In a DAMA system, a separate channel, called the request channel, is used by the individual user to request capacity when it is needed. The way this capacity is allocated is either by the master station in response to the user request or by a common algorithm running in each terminal [35]. Of course, DAMA may be applied to both TDMA and FDMA.

The price paid for accommodating varying traffic by DAMA is the introduction of additional overhead into the multiple access channel due to the process of requesting and (sometimes) assigning capacity [35]. Fortunately, the statistical multiplexing efficiency achieved makes DAMA attractive in the future generation of the personal communications (PCS) [30].

2.3.1 Dynamic Frequency Division Multiple Access (D-FDMA)

Currently, FDMA serves as an auxiliary of TDMA or CDMA to further enhance the system capacity by implementing frequency reuse [30]. In fact, the work of Kohno

[74] can be considered as a D-FDMA on the CDMA platform, where the allocated bandwidth is optimally controlled according to the traffic load and the network performance.

2.3.2 Dynamic Time Division Multiple Access (D-TDMA)

Dynamic TDMA (D-TDMA) has been proposed as a multiple access protocol for satellite and radio communications [12],[13],[25], for local area networks. D-TDMA has different scenarios. For example, in [13], the time frame is divided into three contiguous TDMA frames namely, request slot, voice slot, and data slot. Further, each of these frames are subdivided into TDMA intervals. The boundary between the voice and data slots is adjustable according to the ratio of voice and data traffic. The channel access follows the circuit mode reservation for voice calls, while the data messages are dynamically assigned the available capacity and served according to the first-in-first out policy. The request channel works on a random mode (slotted ALOHA). It is suggested [13] that high D-TDMA capacity may be achievable by employing higher level modulation such as QAM along with trellis modulation.

2.3.3 Packet Reservation Multiple Access (PRMA)

Packet Reservation Multiple Access (PRMA) is another TDMA-based channel assignment multiple access [30] as shown in Fig. 3.4a & b. The PRMA combines random access with time division access and employs voice activity detection to improve the multiplexing efficiency. There is no request channel assigned, but each user (voice or data) desires to transmit should try using the whole information packet. Then, if it succeeds, the user can make reservation. Further, PRMA can support integrated services by assigning voice and data different priorities. The PRMA protocol has been studied in the literature considering several system scenarios [20]. It was shown that PRMA is not efficient and particularly when the traffic is heavy.

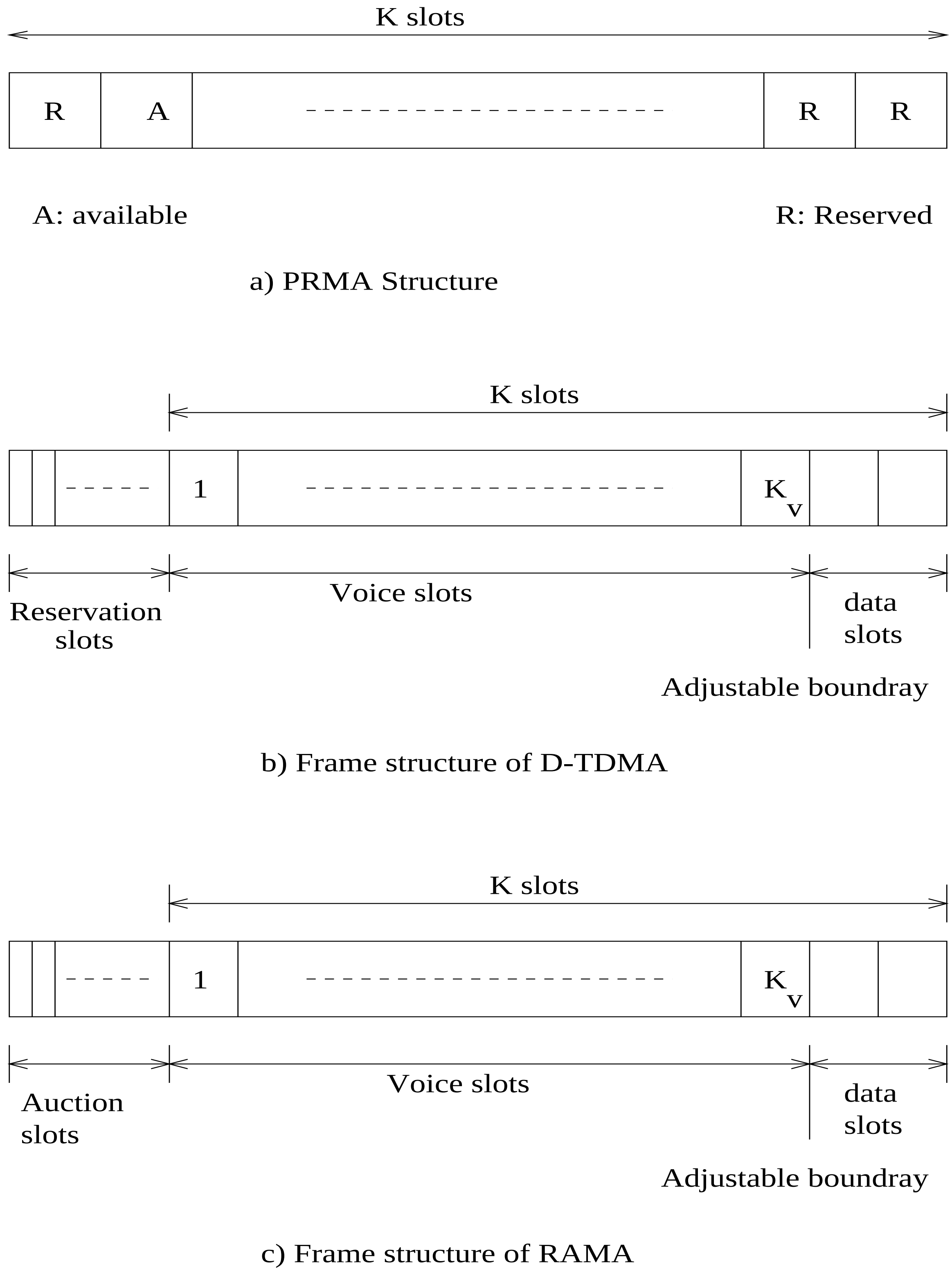


Figure 2.3: Frame Structures of PRMA, D-TDMA, and RAMA

With heaving traffic load, packet collision frequently occurs. PRMA is a slow protocol in the sense that the failure of the user take a duration of one packet to be detected [30].

2.3.4 Resource Auction Multiple Access (RAMA)

The RAMA [36] has the same structure as D-TDMA except that the reservation slot is replaced by an auction slot.Figure 3.4c illustrates the frame structure of RAMA. As was discussed above the D-TDMA uses slotted ALOHA for channel access, while the RAMA uses the auction strategy to access the channel. RAMA works as follows. The unused resources will be "auctioned" to the requesting subscribers and assigned to the winners according to priority. The auction process is based on the user's ID which is a random number generated by the user when he needs to contend [36],[30]. The number of digits in the user'ID depends on the expected number of users in the network, and must be large enough to ensure that the probability of two independent users generating the same ID is very small [30]. Priority digits can be included in the ID to designate the combinations of service priority and fairness considerations [36]. Each unused communications resource will be auctioned one at a time. The subscriber requesting a channel access will transmit its ID, one digit at a time. Following each transmitted digit, the base station announce the highest value among received digits on the downlink channel. All users whose transmitted digit is less than the announced one will drop out from further participation in this assignment cycle. The process is repeated for the next digit and again and again until all the digits have been transmitted, then there will be a final winner at the end of this auction slot. The resource assigned to the winner is announced by the base station following the auction interval. This resource will be removed from the "unused resources" list. It is clear that there still a chance for having identical ID for different users which cause collision, but it does not cause any erasure or failure. This protocol or strategy always allows a requesting user to be admitted to the system irrespective of the traffic demand as contrary to the D-TDMA or PRMA where the system performance degrades very quickly in the heavy traffic. Nevertheless, The RAMA suffers from the implementation complexity and also has the same problem as D-TDMA that is the fixed bandwidth allocated for reservations [30].

2.3.5 Dynamic Reservation Multiple Access (DRMA)

Dynamic Reservation Multiple Access (DRMA) has been proposed [22] as a strategy to overcome the shortcomings of PRMA, D-TDMA, and RAMA. It is dynamic strategy, because the number of reservation slots change and their position in a frame change with time. In fact, DRMA is similar to PRMA in the sense that there is no bandwidth allocated to reservations. In DRMA, each slot may be used either for information transmission or channel reservation minislots depending on necessity. The user needs to transmit and has transmission permission, randomly chooses one of the reservation slots, and transmits a short reservation packet. As most of the protocols proposed to serve integrated services (voice, data, etc.), DRMA gives higher priority to the voice calls. Hence, we shall end with a contention problem that will be solved by the voice transmission permission probability and the data transmission permission probability . At the end of this reservation slot, the assignment is broadcasted by the base station and each successful user is assigned one of the available slots. DMRA shows superiority over the existing demand channel protocols [30]. The literature is rich of other variations and scenarios of DAMA based on TDMA structures e.g [21],[23],[37].

2.4 Random-Access Protocols

What we have mentioned above is the main existing multiple access techniques that modify the basic TDMA frame structure to accommodate the variations in the traffic. Further, the above techniques add another dimension to the TDMA by considering the random access techniques such as ALOHA in the request channel. However, if the traffic is bursty or the required overhead associated with such protocol is comparable to the amount of information to be transmitted, the use of DAMA architecture, even with random access techniques, may not provide an acceptable level of network efficiency [35]. In these circumstances, the only practical alternative is to use the random access technique in the primary channel rather than just in the request channel [35]. There are two general classes of random multiple access techniques: 1) ALOHA and 2) CDMA.

2.4.1 ALOHA Multiple Access

In this protocol, the packets are buffered at each terminal and transmitted over a common channel to the hub station. There is no synchronization between users. Usually, the traffic modeled as a Poisson point process. Under a moderate traffic, there is a high probability of success. However, as the traffic load increase, the probability of collision between different users increases 2,[35].

2.4.2 CDMA Techniques

CDMA techniques are very promising multiple access schemes and their structures make them very competitive candidates for PCS [30]. First, the whole bandwidth is available for all users in each cell, hence, complex frequency planning is avoided that makes the future expansion more easier. Voice activity exploitation and frequency diversity are inherent features of CDMA. Anti-multipath capability is another attractive feature of CDMA. In [13],it is shown that the packet CDMA does achieve a significantly higher bandwidth efficiency than TDMA.

However, CDMA has also some limitations. The performance of DS-CDMA is very sensitive to the accuracy of the power control. Thus a tight power control is necessary for acceptable performance. On the other hand, FH-CDMA does not require very accurate power control, but the required frequency synthesizer is complex. Moreover, to achieve good performance, the data rate should relatively be low, consequently, longer delay will be suffered by long messages such as file transfer. Moreover, packet CDMA system performance is sensitive to the propagation loss constant (γ) [13].

2.5 Hybrid Multiple Access Systems

From the above discussion, it is clear that there is no access scheme that outperforms all others under all conditions [38]. The aim is to accommodate a combination of traffic types. This aim can be achieved by designing adaptive protocols that offer good performance over a wide range of conditions which means that these protocols have flexibility such that the access scheme itself changes, adapting smoothly to network load fluctuations, yielding an access procedure appropriate for the actual state of the network [38]. In other words, we seek from the hybrid system the adaptability to the changes in the system parameters such as traffic dynamics, variations in the required quality of services, etc. In [26],[38], there is another proposed approach to achieve the above objectives that is to partition the channel into several sections, each operating under its own protocol. Here, the focus shall be on those proposals utilizing the CDMA system as one of their components.

2.5.1 Hybrid FDMA/CDMA

FDMA/CDMA was investigated in [11] under Ricean multipath fading where two multipath densities profiles (MIP) for the Rayleigh components of the channel were considered that are: a constant MIP, and an exponential MIP. The hybrid system is formed by dividing the total available bandwidth into separate and isolated 'narrowband' CDMA systems. It is shown that the "wideband" CDMA system outperforms a hybrid F/CDMA. The results also show that under the exponentially decaying multipath MIP with a large decay factor may result in only a minimal degradation of the system performance for a hybrid system.

Jianming and Kohno [74] have investigated a new FDMA/CDMA multiple access protocol for wireless multimedia networks. The total available bandwidth is divided into a number of sub-band frequency. Each medium is assigned one of these sub-bands. The width of each sub-band is optimally controlled in terms of the processing gain which is obtained by using nonlinear programming. The control factors are the traffic dynamics and the prescribed priority for each medium. This system has shown higher flexible capacity than TDMA in a dynamic multi-media traffic channel. Nevertheless, the receiver structure is expected to be too complicated because it is required to be adaptive to the dynamic changes in the bandwidth of each sub-band.

2.5.2 Hybrid CDMA/TDMA

Elhakeem et al [12] analyzed the delay and throughput performance of fixed frame TDMA/CDMA, variable frame TDMA/CDMA, and adaptive time hopping CDMA for both voice and data transmission under Rayleigh and Ricean multipath fading channels. It is shown that variable frame TDMA/CDMA and adaptive time hopping CDMA provide the best delay results for voice traffic in both Rayleigh and Ricean fading channels and nonfading channels.

In [34] a CDMA/TDMA mobile system applying joint detection and coherent receiver antenna diversity has been proposed. In fact this proposed protocol uses the same TDMA structure used in the GSM mobile system with the exception that each user spreads his own signals using DS-CDMA. Further, a forward error correcting code is used to mitigate the multiple access interference (MAI) effects.

2.5.3 Hybrid FDMA/TDMA/CDMA

This multiple access technique was proposed [15] as an air interface for Universal Mobile Telecommunication system (UMTS). The different advantages of each multiple access scheme are combined to obtain the utmost performance.

2.6 Time Division Duplex (TDD)

In the time division duplex system (TDD) [14], the same frequency band is used by both links (i.e uplink and downlink). TDD system has some advantages over the frequency division duplex when there is no balance in data rate between the uplink and the downlink. The FDD system has to have a quite narrow filter to distinguish a low rate signal that is transmitted among high speed signal. Moreover, a frequency synthesizer is required to choose an RF carrier frequency that wastes a valuable battery power. On the other hand, time synchronization is essential in the TDD system. However, in TDD, the channel impulse response can easily be estimated by using the CDMA signal transmitted in a TDD time slot, where the MS estimates the forward channel impulse response and the BS estimates the reverse channel impulse response. Then, both parties exchange these information. These measurements can be used to control the transmitted power in the uplink and / or the channel equalization in the down link. In [14],[31],[32], TDD-CDMA/TDMA multiple access protocol is investigated where the radio channel for the proposed mobile network is assumed an indoor micro cellular channel. The uplink employs a CDMA scheme to transmit low speed human interface signal, while the downlink employs a TDMA scheme to transmit high speed video signals

2.7 Conclusions

A brief review of the existing multiple access protocols has been presented. No protocol can be claimed that it outperforms all others under all conditions. Therefore, Hybrid access schemes have also been explored. The comparison between these access scheme should be done taking in consideration the traffic characteristics, performance measures, the existing technology at the time of deployment of such network.

Chapter 3

CDMA Receivers

3.1 Introduction

The structure of the CDMA signals is quite different from the structures of the other orthogonal multiple access techniques such as FDMA and TDMA. This structure has its unique features as well as its unique problems. For example, it is well known that CDMA communication systems are interference limited. These interferences are due to other users who are sharing the same frequency channel simultaneously. The detection process in such environment is not straight forward, because the desired signal is corrupted by other signals which are intended by other users. Therefore, the interference consists of two main parts: the interferences due to other users and the additive white Gaussian noise. This interesting situation motivates the researchers to propose several scenarios for the CDMA receiver. The proposed receivers can be classified into two classes: *Single-user* receiver (Conventional) and *Multi-user* receiver.

The above classification refers to the way the receiver should treat or deal with undesired signals. The later jointly detects the desired signals using the fact that the receiver already has information about the transmitted signals such as the signature code of each user or it can estimate important characteristics of the transmitted signals (such as energy of the received signals) that can help in improving the process of jointly detecting these signals.

On the other hand, the conventional single user receiver treats the received signal as if it is composed of the desired signal plus noise (e.g. Gaussian Noise). Therefore, this receiver deals with each user separately [42].

In the following sections, we will review in detail both types of receivers and highlight the limitations and potentials of each receiver. To begin with, we repeat the formulation of the DS-CDMA signal and channel model. The receiver front end observes the following composite signal assuming there are K users sharing a white

Gaussian channel.

$$y(t) = \sum_{k=1}^{K} s_k(t - \tau_k) + n(t);\ 0 < t < T_b \tag{3.1}$$

where $s_k(t)$ is the equivalent lowpass transmitted signal of a block of N bits of the $k-th$ user, and τ_k is the transmission delay, which satisfies the condition $0 \leq \tau_k \leq T_b$ for all $1 \leq k \leq K$, T_b being the bit duration assumed to be equal for all users. $n(t)$ is additive white Gaussian noise with power spectral density $N_o/2$.

$$s_k(t) = \sqrt{E_k} \sum_{i=1}^{N} b_k(i) a_k(t - iT_b) \tag{3.2}$$

where E_k is the signal energy per bit. This is the general model for the multiuser transmission. $b_k(t)$ and $a_k(t)$ are the binary phase shift keying (BPSK) baseband data signal and the spreading waveform for the kth user, respectively.

3.2 The Conventional Single User Detector

We shall start with the conventional detector which it is a suboptimal receiver from the point view of the nature of the interferences in the DS-CDMA systems. The conventional or single user detector is the simplest receiver. The interesting thing that stimulates the researchers to use the conventional detector as a receiver for the DS-CDMA systems is that it is well known that the Gaussian noise is the worst kind of additive noise from the capacity standpoint [45],[49],[52]. Further, it is well known [51] that the matched filter is the optimum receiver if the interferences are confined to be a White Gaussian Noise.

Fig. 3.1 illustrates this detector. Simply, the receiver is composed of a bank of matched filters for each user. The received signal is correlated with the signature code associated for each user, then it passes to a detector where a decision should be taken based on the correlator output. Thus, this type of receiver ignores the presence of other users and assumes that any corruption in the received signal is due to an AWGN multiple access channel.

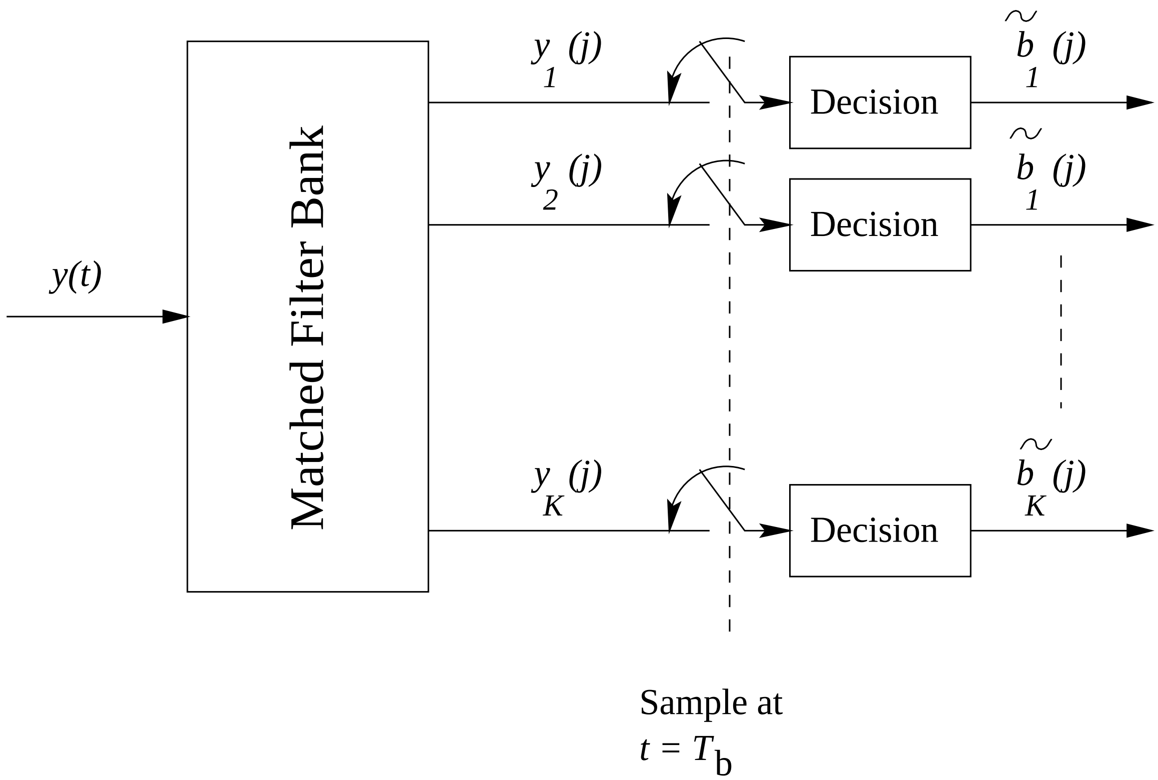


Figure 3.1: The conventional DS-CDMA detector

Assume synchronous transmission, $\tau_k = 0$ for all $1 \leq k \leq K$. The output of the kth correlator is

$$y_k = \int_0^{T_b} y(t) a_k(t) dt \tag{3.3}$$

which can be decomposed into three components: 1) the desired signal, 2) multiple access interference (MAI) from coexisting users, and 3) noise.

$$y_k = \sqrt{E_k} b_k + \sum_{\substack{i=1 \\ i \neq k}}^{K} \sqrt{E_i} b_i \rho_{i,k}(0) + \int_0^{T_b} a_k(t) n(t) dt \tag{3.4}$$

$$\rho_{i,k}(0) = \int_0^{T_b} a_i(t) a_k(t) dt \tag{3.5}$$

where $\rho_{k,k} = 1$. If we examine the output of the kth (i.e. y_k, $1 \leq k \leq K$) correlator, we can observe that the success of this type of receiver is highly dependent on the cross-correlations between codes as it is clear from Eq. 3.4. Theoretically, this interference problem can be overcome by designing orthogonal signature codes.

Unfortunately, Asynchronous transmission and Multipath Fading make it impossible to design signature sequences for any pair of users that are orthogonal for all time offsets [8]. Thus, interference from other users is unavoidable and the larger the number of overlapping users, the more the conventional detector is vulnerable to the multiple access interference. On the other hand, in practice, the signals received by the receiver in the base station suffer different levels of attenuation due to two factors: Fading and mobile's geographical location (far or near) relative to the base station. Consequently, users near the receiver end are received at high powers and those that are far away are received at low powers. This is called near-far effect [42],[8]. Thus, even if we could design signature codes that have very low cross-correlations relative to autocorrelations (i.e. $\rho_{k,k} \ll 1$), the diversity of the strength of the received signals will have a significant impact on the capacity and performance of the conventional detector [43]. This introduces the necessity for a tight power control [42],[43] to achieve a satisfactory performance.

Now, we turn to examine the performance of the conventional detector. Here, we shall concentrate on the error rate of this type of detector. As mentioned above, the conventional DS-CDMA detector follows the same approach of a single user detector where the error rate is upper bounded by

$$P_e(k) = Q\left(\sqrt{SNR^k_{eff}}\right) \tag{3.6}$$

where $P_e(k)$ is the probability of error performance of the kth, SNR^k_{eff} is the effective signal-to-noise ratio of the k-th user which is defined to be E_k/I_o. I_o is the total interference power density, E_k is the signal energy per bit of the k th user and Q-function is defined as

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-y^2/2} dy \tag{3.7}$$

The question raising now is how to evaluate I_o in the case of DS-CDMA? Modeling the multiple access interferers as random Gaussian variables [53]-[55] leads to the following formulation of the effective signal-to-noise ratio

$$SNR_{eff}^{k} = \left\{ (E_k/N_o)^{-1} + \frac{K-1}{3PG} \right\}^{-1} \tag{3.8}$$

where PG is the number of chips spanned in one bit period (i.e. $PG = \frac{T_b}{T_c}$). We end this section by the following observations. First, the structure of the conventional detector is very simple and its performance is well understood. Yet, the error rate under practical considerations (i.e. SNR, K, PG, etc.) is prohibitive. There are several ways to mitigate the degradation in performance, namely: Error control coding 17, Long signature codes with very low cross-correlations [56], power control [99]-[61] to ensure fairly strength of received signal at the receiver end. In the following sections, we shall present other detectors proposed in the literature as alternative candidates to DS-CDMA systems.

3.3 The Optimum Multi-User Detector

The optimum multiuser detection is defined as the strategy that selects the most probable transmitted sequence of bits given that the observation is done over the whole received waveform. This is required to produce a sufficient statistics for any bit decision. Having observed the whole waveform, an additional information is obtained about the received signal in the bit interval under consideration. Hence, the single-user approach is suboptimal [41],[8]. Here, we shall present the analysis of the multiuser detectors under the asynchronous transmission. For more details, [8], [41] and [44] are recommended.

3.3.1 Asynchronous Transmission

Considering the uplink transmission (Mobile-to-Base station, or Earth station to Satellite), where each user can transmit at any time, the asynchronous scheme of transmission is the reality of such systems. On the other hand, considering the

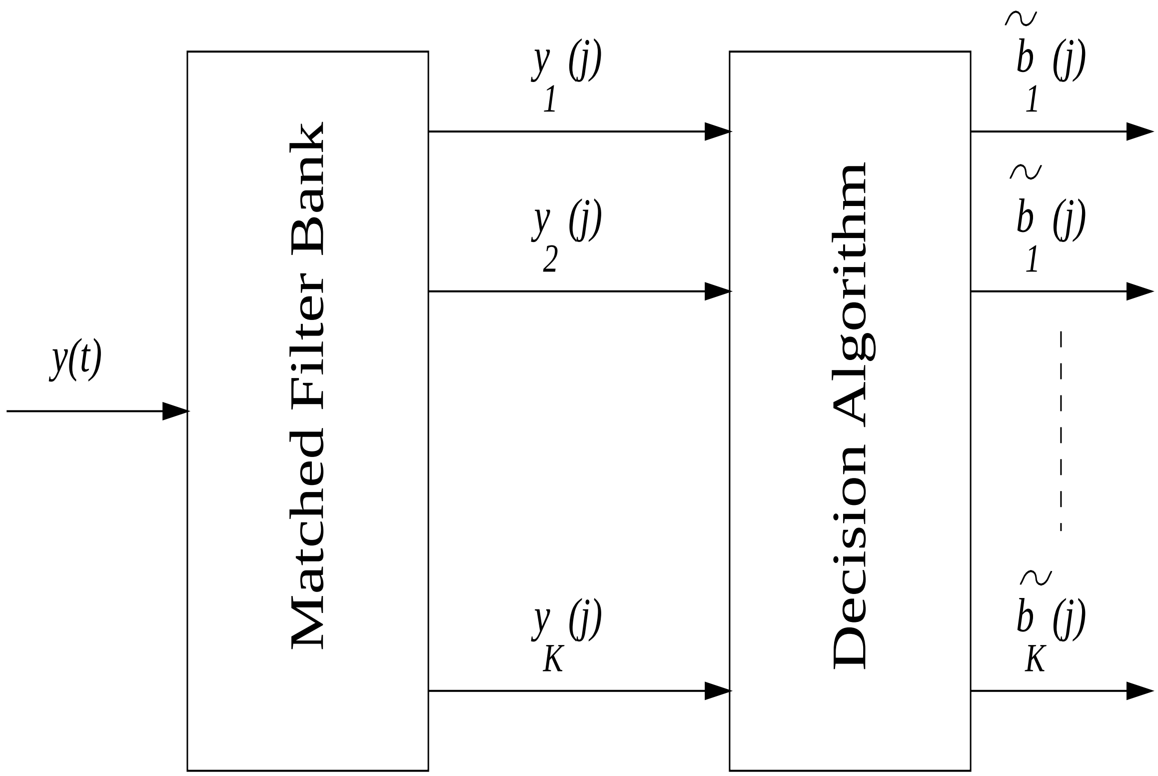


Figure 3.2: A general multiuser DS-CDMA detector

down link, the transmission can easily be synchronized and each interfere produces exactly one symbol that interferes with the desired symbol. Using the same notation as above, and assume that we just have an additive white Gaussian noise, then the optimum maximum-likelihood receiver computes all users bit sequences that minimizes the following log-likelihood function

$$\begin{aligned}
\Lambda(b) &= \int_0^{NT_b+2T_b} \left[y(t) - \sum_{k=1}^{K} \sqrt{E_k} \sum_{i=1}^{N} b_k(i) a_k(t - iT_b - \tau_k) \right]^2 dt \\
&= \int_0^{NT_b+2T_b} y^2(t)\, dt - \sum_{k=1}^{K} \sqrt{E_k} \sum_{i=1}^{N} b_k(i) \int_0^{NT+2T_b} y(t) a_k(t - iT_b - \tau_k)\, dt \\
&+ \sum_{k=1}^{K} \sum_{l=1}^{K} \sqrt{E_k E_l} \sum_{i=1}^{N} \sum_{j=1}^{N} \int_0^{NT_b+2T_b} a_k(t - iT_b - \tau_k) a_k(t - jT_b - \tau_l)\, dt \qquad (3.9)
\end{aligned}$$

Examining Eq. (3.9), we notice that the first term involves $y^2(t)$ which is common for all possible data sequences, and hence it can be ignored. The second term involves the integral

$$y_k(i) \equiv \int_{iT_b+\tau_k}^{(i+1)T_b+\tau_k} y(t)a_k(t - iT_b - \tau_k)\, dt \quad ,1 \leq i \leq N \tag{3.10}$$

which represents the correlator or matched filter outputs for the kth user in the observation interval as shown in Fig. 3.2. From Eq. 3.1 and Eq. (3.10), we observe that the matched filter outputs $\{y_k(i)\}$ can be expressed as follows.

$$\begin{aligned} y_k(i) &= \sum_{m=1}^{K} \sqrt{E_m} \sum_{j=1}^{N} \int a_k(t - iT_b - \tau_k)a_k(t - jT_b - \tau_m)\, dt \\ &\quad + \int n(t)a_k(t - iT_b - \tau_k)\, dt \end{aligned} \tag{3.11}$$

It is clear that the integral in the first term in Eq. (3.11) represents the cross-correlation between the signature waveforms. Let $\rho_{km}(i-j)$ be the cross-correlation between the signature waveforms of user i and j,

$$\rho_{km}(i-j) = \int a_k(t - iT_b - \tau_k)a_k(t - jT_b - \tau_m)\, dt \tag{3.12}$$

where $\rho_{km}(i-j) = 0,\ |i-j| > 1$ are the entries of the KxK cross-correlation matrix $R(\ )$. Now, it is easy to express the matched filter outputs in the vector notation as shown in Eq. (3.13).

$$\mathcal{Y} = \mathcal{R}\ \mathcal{B} + \mathcal{N} \tag{3.13}$$

where by definition,

$$\mathcal{Y} = [Y(1)\ Y(2) \cdots\ Y(N)]^t, \quad r(i) = [y_1(i)\ y_2(i) \cdots\ y_K(i)]$$

$$\mathcal{B} = [B(1)\ B(2) \cdots\ B(N)]^t, \quad B(i) = \left[\sqrt{E_1}b_1(i)\ \sqrt{E_2}b_2(i) \cdots\ \sqrt{E_K}b_K(i)\right]$$

$$\mathcal{N} = [\mathbf{n}(1)\ \mathbf{n}(2) \cdots\ \mathbf{n}(N)]^t, \quad \mathbf{n}(i) = [n_1(i)\ n_2(i) \cdots\ n_K(i)]$$

$$\mathcal{R} = \begin{bmatrix} R(0) & R^t(1) & 0 & \cdots & 0 & \cdots & 0 \\ R(1) & R(0) & R^t(1) & 0 & \cdots & \cdots & 0 \\ 0 & R(1) & R(0) & R^t(1) & 0 & \cdots & 0 \\ 0 & \cdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & R(1) & R(0) & R^t(1) & 0 \\ \vdots & \vdots & \vdots & \vdots & R(1) & R(0) & R^t(1) \\ 0 & 0 & 0 & \cdots & \cdots & R(1) & R(0) \end{bmatrix}$$

Such an approach leads to the following conclusions. First, the receiver should have the ability to estimate both the received energy of each arrived symbols $\{E_k\}$ and the transmission delays $\{\tau_k\}$. Second, we observe the huge dimension $\{NK\text{x}NK\}$ of the cross-correlation matrix that is required to be computed for each decision. Furthermore, the optimum receiver is expected to compute this matrix for all possible transmitted sequence. Therefore, the computation complexity of such detector is of $O(2^{NK})$. Fortunately, in the case of asynchronous transmission, the desired symbol is exactly overlapped with two consecutive symbols from each interferer, providing that all users have the same symbol rate. Therefore, each transmitted symbol overlaps at most with $(2K-2)$ symbols. Using this fact makes it possible to use the sequential detection employing the Viterbi algorithm. Hence, the computational complexity is reduced considerably to be of $O(2^K)$. Nevertheless, the exponential dependence on K can not be reduced. This motivates the researchers to look for sub-optimal detectors with a computational complexity of linear dependence on K.

3.4 The Sub-Optimum Multi-user Detectors

The sub-optimum detectors can be classified into two groups, namely, Linear and Nonlinear detectors. This classification is done according to the way the matched

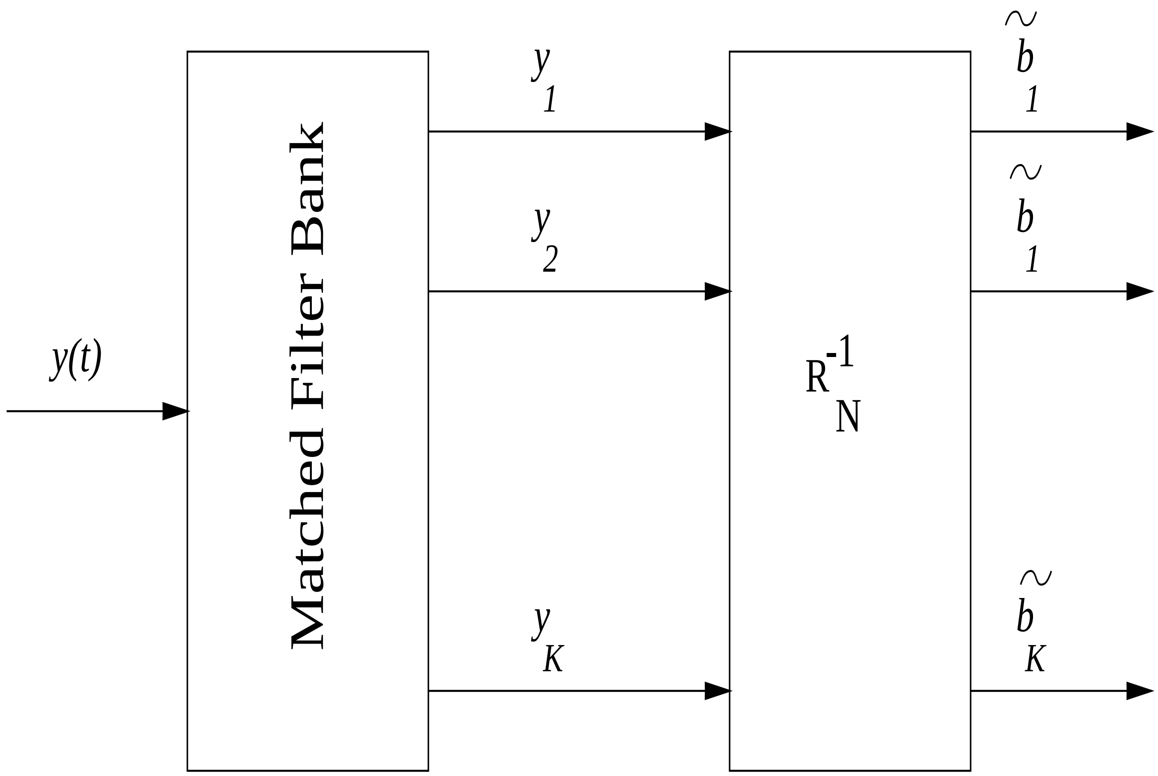


Figure 3.3: The decorrelator multiuser DS-CDMA detector

filter outputs are transformed such that the transmitted bits are estimated.

3.4.1 Linear Multiuser Detectors

Decorrelating Detector

We observed the exhaustive search required in the maximum-likelihood detector for each symbol decision. However, the decorrelator receiver uses a linear transformation to obtain an estimate of the transmitted symbol. From (3.13), we observe that the solution (i.e. the bit vector $\mathcal{B}_K$) can be obtained by inverting the correlation matrix $\mathcal{R}_K$ which is invertible in most cases of interest [62]. Hence,

$$\mathcal{L} = \mathcal{R}_K^{-1} \tag{3.14}$$

This results in

$$\widetilde{\mathcal{B}}_K = \mathcal{B}_K + \mathcal{R}_K^{-1}\mathcal{N}_K \tag{3.15}$$

where $\widetilde{\mathcal{B}}_K$ is estimated data vector $\left[\tilde{b}_1\tilde{b}_2\cdots\tilde{b}_K\right]^T$.

Figure 3.3 depicts the decorrelator multiuser structure. Examining Eq. (3.15), we can extract many attractive features for this multiuser detector. First of all, the multiuser interference (MAI) has been eliminated. Second, its computational complexity is much reduced comparing with the optimum receiver. The computational complexity of the decorrelator detector increases linearly with the number of users (i.e. $O(K)$). Third, it uses a linear transformation to obtain an estimate of the transmitted symbol. At practical SNR values, the decorrelator receiver provides much better performance than the conventional one. Further, the linear decorrelator receiver exhibits the same degree of near-far resistance as the optimum multiuser detector. In addition, when the users energies are unknown, the decorrelator receiver is the optimal approach [62].

Now, we consider the performance characteristics of the synchronous decorrelator receiver under AWGN. Therefore, $\tau_k = 0$ for all $1 \leq k \leq K$. The output of the kth correlator is

$$y_k = \int_0^{T_b} y(t)a_k(t)dt \tag{3.16}$$

$$\rho_{i,k}(0) = \int_0^{T_b} a_i(t)a_k(t)dt \tag{3.17}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \\ \vdots \\ y_K \end{bmatrix} = \begin{bmatrix} 1 & \rho_{1,2} & \cdots & \rho_{1,k} & \cdots & \rho_{1,K} \\ \rho_{2,1} & 1 & \cdots & \rho_{2,k} & \cdots & \rho_{2,k} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \rho_{k,1} & \rho_{k,2} & \cdots & 1 & \cdots & \rho_{k,K} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{K,1} & \rho_{K,2} & \cdots & \rho_{K,k} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \sqrt{E_1}b_1 \\ \sqrt{E_2}b_2 \\ \vdots \\ \sqrt{E_k}b_k \\ \vdots \\ \sqrt{E_K}b_K \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_k \\ \vdots \\ n_K \end{bmatrix} \tag{3.18}$$

where

$$n_k = \int_0^{T_b} a_k(t)n(t)dt \tag{3.19}$$

Now, we can show [8] that the probability of error is given by

$$P_e(k) = Q\left(\sqrt{\frac{E_k}{\mathcal{R}_{k,k}^{-1} N_o}}\right) \tag{3.20}$$

where $P_e(k)$ is the probability of error performance of the kth, $\mathcal{R}_{k,k}^{-1}$ is the (k,k) entry of inverse of the correlation matrix, N_o is the one-sided power spectral density of AWGN, E_k is the signal energy per bit of the k th user and Q-function is defined as

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-y^2/2} dy \tag{3.21}$$

However, a significant limitation of this technique is the computational complexity due to the inversion of the correlation matrix [45] whose entries depend on the number of active users, signature sequences, and the users delays. Many researches have tried minimizing the computations required, for example Ref. [45]. Further, any change in one of these parameters changes the correlation matrix and consequently a need for updating the multiuser detection process. Moreover, the uncertainty in the actual number of active users is another serious problem that might degrade the system performance very severely [66] which could now be resolved by applying our proposed scheme for traffic control.

MMSE Detector [8][43]

The minimum mean squared error detector is a linear detector which balances between the desire to decouple the desired user from the other users interference and at the same time avoids the enhancement of the noise power as in the decorrelator detector by utilizing the knowledge of the received signal powers. The MMSE detector is exactly analogous to the MMSE equalizer used to combat ISI. It is obvious that the main disadvantage of this detector is its dependence on the amplitude estimation of the received signals. Further, it suffers from the same problem as the decorrelator detector that it needs a matrix inversion.

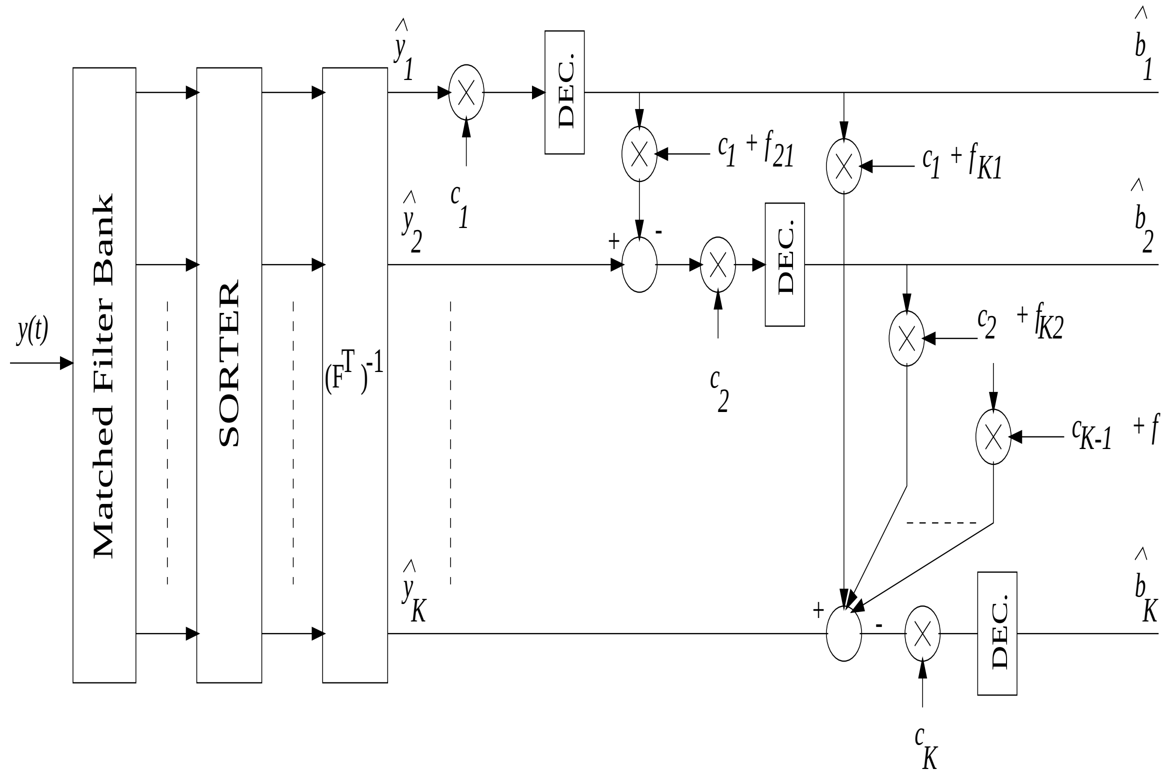


Figure 3.4: The decision feedback detector for synchronous CDMA

Polynomial Expansion (PE) Detector [43]

The PE detector approximates the decorrelator receiver without the need for the matrix inversion. It is a linear detector that applies a linear transformation extracted from the cross-correlation matrix $\mathcal{R}$. The structure of the PE detector is composed of several stages following the conventional detector. Each stage is simply a spreader (modulator) and a matched filter bank (demodulator). The outputs of each stage including the conventional receiver stage are weighted and summed together. The weights can be chosen to optimize some performance measure.

3.4.2 Non-Linear Multiuser Detectors

[43] The linear detectors mentioned above have provided significant improvements over the conventional detector. Nevertheless, their linear structure often limits their

performance. Generally, the nonlinear structured receivers outperform the linear counterparts.

The basic principle behind the nonlinear detector can be clarified by the following example. Let us have two users communicating with the base station. At the receiver front-end, we would receive the following baseband signals (note that the notations used here are the same as above):

$$y_1 = \sqrt{E_1} b_1 + \rho_{1,2} \sqrt{E_2} b_2 \tag{3.22}$$

$$y_2 = \sqrt{E_2} b_2 + \rho_{2,1} \sqrt{E_1} b_1 \tag{3.23}$$

Assume that user1 is much stronger than user2 (i.e. near-far problem), then the decision of the conventional receiver is reliable. Now we can use the decision on user1 to cancel its effect on user2 as shown below in Eq. (3.24).

$$\begin{aligned} \widetilde{b}_2 &= \mathrm{sgn}\{y_2 - \rho_{2,1}\sqrt{E_1} b_1\} \\ &= \mathrm{sgn}\{\sqrt{E_2}\widetilde{b}_2 - \rho_{2,1}\sqrt{E_1}(b_1 - \widetilde{b}_1)\} \end{aligned} \tag{3.24}$$

In the literature, we can find several proposed nonlinear multiuser detectors that implement the above principle. For example, in [63] and [64], the multi-stage detector was proposed, and a family of decision-feedback detectors were introduced in [65]. In general, the nonlinear detectors follow certain steps in the process of decision making. The main common steps are 1) Estimation of the received signals amplitudes, 2) Regeneration and 3) Cancellation. Further, some nonlinear detectors have an additional step that is sorting the received signals in an ascending order according to their energies (e.g. decision-feedback detectors) to cancel stronger interferences first and limit the number of stages.

Each of the above nonlinear detectors has an initial stage as well as successive stages (see Fig. 3.4 as an example). The initial stage provides the detector with initial guess regarding the transmitted signals. This initial stage would be a conventional receiver or decorrelator receiver. Therefore, the improvement in the performance obtained by such detector is highly dependent on the decisions taken in the initial stage. Moreover, in the regeneration process the receiver tries to create signals resemble the transmitted ones. It is obvious that this step is very crucial and the improvement of such detector is highly dependent on the accuracy of the regeneration step, because the cancellation step presume that the information obtained is correct.

3.4.3 Implementation Issues

The multiuser detection strategy has very attractive features that makes it promising candidate for CDMA communications and personal communications networks. Nonetheless, it is still a long way before this strategy overcomes the practical implementation issues. In fact, there are several problems that need to be investigated more such as the computational complexity, the tracking error (frequency, amplitude, phase, and timing uncertainties) [42], and most of all the need to accurately know the actual number of users packets at any time. In fact, the last issue is very crucial in the performance improvement obtained by the multiuser detectors. Recently, Esteves and Scholtz [66] show the severe degradation in the performance of some multiuser detectors when the exact number of active users is unknown.

Chapter 4

Admission/Congestion Control Policies for CDMA Networks

4.1 Introduction

With the introduction of multimedia services for the third generation wireless networks has faced the researchers with many new challenges that should be resolved in order to deploy reliable end-to-end service networks. One of these problems is the issue of control and management of traffic flow into the wireless network. This is very important issue for properly designing and operating CDMA networks.

Since covering all related aspects of admission control is beyond the scope of this work, we are going to concentrate in our survey on the work done related to CDMA integrated networks. However, interested readers could find a wider survey in [67]-[70]. The decision to accept or reject a call is based on the following questions [67]. First, does the new call affect the QoS of active callers currently carried by the network? Second, can the network provide the QoS required by the new caller? The question then becomes how can we map the end-to-end QoS requirements into physical parameters. In other words, how can we map the link layer QoS into the physical layer QoS [76]. How does the network admit new calls providing that certain quality of service requirements are guaranteed? How can we estimate or predict the network performance if a new call is admitted?

Recently, researchers have become very active in investigating this problem and finding solutions for new issues associated with it. Typically, admission/congestion control can be classified into *preventive* control and *reactive* control. In *preventive* congestion control, the schemes are designed to prevent the occurrence of congestion, while in the *reactive* congestion control, one relies on the feedback information for controlling the level of congestion [67]. Further, looking at the matter from the prospective on which the decision is made, we can categorize the existing admission decision strategies into two main classes. First, strategies where the figure of merit is the signal-to-interference ratio (**SIR**) such that each class of traffic is guaranteed that its BER requirement is met. These admission strategies are evaluated according to their effectiveness in increasing the system overall capacity and decreasing the outage probability which is defined as the percentage of time that (**SIR**) is less than the minimum (**SIR**) requirement. The second category of admission decision strategies is based on traffic activity at the base station.

4.2 Power Assignment Problem

Let us assume a single cell CDMA system where there are N active users belonging to the same class of traffic and the QoS parameter is the same for all (e.g. BER $\leq 10^{-3}$). Therefore, the maximum capacity is achieved when all users (presumably independent) could be controlled such that they arrived at the BS front with the same bit-energy (i.e. E_b/I_o) level [28]. Hence, the demodulator processes a composite received waveform containing the desired signal having power S and $(N-1)$ interfering signals each also with power S. Now, the signal-to-interference ratio (SIR) which is defined as the ratio of signal power to the total noise (mutual interference as well as background noise) power, i.e.

$$\text{SNR} = \frac{S}{(N-1)S + \eta} \tag{4.1}$$

where η is the background noise power. It is of greater importance to find the bit energy-to-noise density ratio (i.e E_b/I_o). From Eq. (4.1), divide the numerator by the information bit rate , R, and divide the noise level by the total bandwidth, W, we get the following expression:

$$\begin{aligned} E_b/I_o &= \frac{S/R}{(N-1)S/W + \eta/W} \\ &= \frac{W/R}{(N-1) + N_o/S} \end{aligned} \tag{4.2}$$

Equation (4.3) implies that the capacity in terms of number of supported users is

$$N = 1 + \frac{W/R}{E_b/I_o} - \frac{\eta}{S} \tag{4.3}$$

where W/R is usually referred to as the processing gain. Ignoring the background noise, It is clear from Eq. (4.3) that the system capacity is inversely proportional to the required E_b/I_o per user. Therefore, CDMA system's capacity is not fixed and it is essential to reduce the per user transmitted power to achieve high system capacity. This goal can be achieved via different means such as, 1), using powerful forward error correction (FEC) codes such as concatenated RS codes or product Turbo codes, 2) applying cell sectorization technique where directional antennas are used at the cell site both for receiving and transmitting, 3) taking advantage of the packet-level users activity. Nevertheless, we should note that the performance of CDMA networks is very dependent on users' locations and propagation parameters. Hence, the CDMA network could not have full advantages of all these mentioned techniques without controlling the user transmitted power [71].

4.2.1 Multiple Cell System

Figure 5.13 illustrates an imaginary hexagonal cell boundaries of a cellular system. Suppose there is a mobile user at distance r_h from his home base station, then the average received signal strength $\Gamma(r)$ in real value can be expressed as:

$$\Gamma(r) = S10^{\zeta/10}r^{-\alpha} \tag{4.4}$$

where ζ in decibels has a normal distribution with zero mean and standard deviation of σ which is independent of the distance and ranges from 5-12 dB with a typical value of 8 dB. Further, due to the existence of other users in the network communicating via other base stations, the interference produced would be very large and it should be considered in the power control process. Let us take a simple example where a mobile user i at a distance r_h from his home base station and at a distance r_k from base station k. So, this mobile user shall interfere with users at BS_k with the following noise component:

$$I(h,k) = S\left({r_{ik}}^{-\alpha}10^{\zeta_{ik}/10}\right)\left({r_{ih}}^{-\alpha}10^{\zeta_{ih}/10}\right)^{-1} \tag{4.5}$$

where the first term is due to the attenuation and blockage to the desired cell (i.e. k) and the second term is the effect of power control to compensate for the

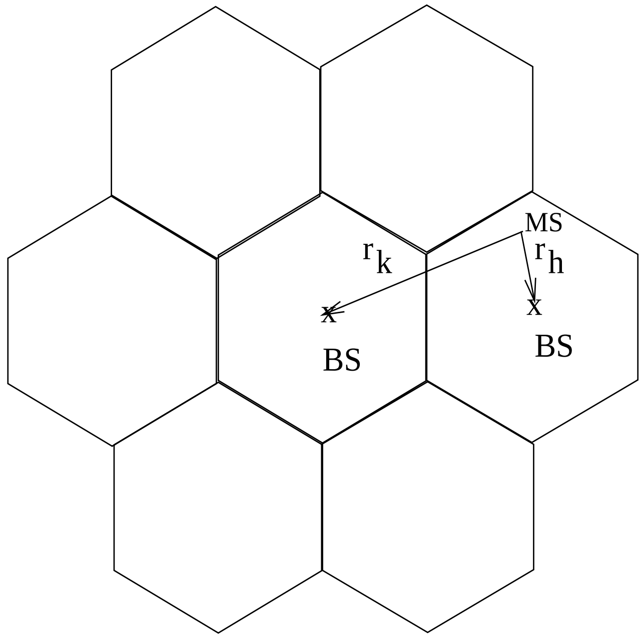


Figure 4.1: Imaginary hexagonal cell boundaries of a cellular system

corresponding cell site of the out-of-cell interference. ζ_{ih} and ζ_{ik} are independent random variables with normal distribution with zero mean and standard deviation σ_{ih} and σ_{ik}, respectively. Consequently, the total out-of-cell interference can be expressed as follows assuming M cells:

$$\begin{aligned} I(k) &= S \sum_{\substack{h=1 \\ h \neq k}}^{M} \sum_{i=1}^{n_h} I_i(h,k) \\ &= S \sum_{\substack{h=1 \\ h \neq k}}^{M} \sum_{i=1}^{n_h} \left(\frac{r_{ih}}{r_{ik}} \right)^{\alpha} 10^{(\zeta_{ik} - \zeta_{ih})/10} \end{aligned} \tag{4.6}$$

where n_h is the number of active users in cell h. Therefore, it is essential to control the received power such that certain quality of service (generally BER) is guaranteed. Hence, the capacity of these networks is not fixed.

Now, the problem is how can the base station allocate to each user what it needs such that the total transmitted from all users is minimized? Further, the problem becomes more complicated when the traffic population includes diverse classes of users where each class requires different QoS parameters.

4.2.2 Current CDMA System

In a CDMA system, the cell site continually measures the received signal from the mobile, compares it to the desired power level, and then makes a decision to raise or lower a specific mobile's transmit power as frequently as once every 1.25 milliseconds (800 times per second). At certain locations, the signal received by a mobile unit may be too weak to accurately decode data (excessive shadowing, interference from neighboring cell–signals from neighboring cells don't experience same degree of fading as signals from the mobile unit's own cell). However, transmission power should only increase when it is necessary. The forward links power control mechanism of the IS 95 standard maintains that the cell periodically reduces the transmitted power. When a mobile detects an increase in its frame error rate, it requests higher power. Then, cell site increases power by a predetermined amount (.5 dB) once per vocoder frame or about 15-20 msec. Further, the dynamic range is limited to +/- 6 dB, so as to set that power level as low as possible while still maintaining a high quality call. Any unneeded power adds unnecessarily to the overall noise level on the CDMA channel, and cuts down capacity. Therefore, the more precise the power control, the greater is the capacity.

The literature has numerous researches on algorithms focusing on how to control and allocate power for active users without associating the admission/congestion issues. This is not the scope of this work and it will not be discussed later. However, what we are interested with in this work is the joint admission/congestion control policy with power control to guarantee certain QoS parameters. In other words, we shall concentrate on those schemes that link between power control issue and accepting/rejecting new users to the network. The existing work on this important issue can be classified into three classes: 1) Power control-based policies, 2) Rate control-based policies and 3) Combined power/rate control policies.

4.3 Power Control Based Admission policies

The philosophy behind this approach is that the source which should be managed and allocated in CDMA system is *Energy*. Hence, this scheme of admission/congestion control depends on using the signal-to-interference ratio (SIR) measured at the base station of cell k, SIR_k to decide whether a new call could be accepted given that certain service quality is maintained.

The work in [79] was one of the earliest in this direction, where the notion of residual capacity $\mathcal{R}_k$ is introduced and defined as the additional number of initial call the base station can accept such that the system-wide outage probability, defined as the probability that an acceptable transmission quality can not be maintained, will be guaranteed to be below a certain level. $\mathcal{R}_k$ is defined in terms of measured SIR_k at the base station receiver, i.e.

$$\mathcal{R}_k = \begin{cases} \lfloor \frac{1}{\text{SIR}_{TH}} - \frac{1}{\text{SIR}_k} \rfloor & ; \lfloor \frac{1}{\text{SIR}_{TH}} - \frac{1}{\text{SIR}_k} \rfloor > 0 \\ \\ 0 & ; \text{otherwise} \end{cases} \tag{4.7}$$

where $\text{SIR}_{TH} > \text{SIR}_O$, and SIR_O is the minimum **SIR** for the proper operation. For each initial call request, the base station checks the value of the residual capacity $\mathcal{R}_k$: if $\mathcal{R}_k > 0$, the new call is accepted and the residual capacity is reduced by one: otherwise, the call request is rejected. Two traffic admission algorithms are compared under nonuniform traffic conditions. In the first algorithm, the call admission decision is solely based on the **SIR** measurement at the local cell base station and ignoring the out-of-cell interference. The second algorithm, the admission decision is based on the local **SIR** measurement as well as the **SIR** measurements of the neighboring cells. In both algorithms, it is assumed that the SIR's are known to all base stations.

In [73], an adaptive SIR-based call admission control is investigated. The decision is based on estimation of the interference ($\hat{I}$) at the base station receivers using a linear Kalman filter which is driven by a measurement of the interference and predicted traffic parameters (i.e. mean (i_p) and variance (v_p)) caused by the accepted connections. the final admission decision is based on the following criterion.

$$\hat{I} + i_p \leq I_{max} - r \tag{4.8}$$

where I_{max} is the maximum interference threshold which ensures that the **BER** constraint is met, r is a reservation threshold. The proposed strategy was tested considering both local and global interferences. the interesting conclusion was that the SIR based only on local information can provide adequate performance. However, providing the scheme with global information, the overall performance could be improved up to 15% depending on the propagation factor and environment variability.

The essence of the work in [77] is similar to those intended by the SIR-based admission control that the QoS parameter to be guaranteed is the *packet error rate*; $P_E(k)$. Defining K_v and K_d as the number of simultaneous voice users and data users, respectively that can be served so that the expected packet error probabilities remain below known thresholds:

$$P_E(k) \leq P_E^v \quad \forall\, k \leq K_v$$

$$P_E(k) \leq P_E^d \quad \forall\, k \leq K_d$$

where P_E^v and P_E^d are the maximum tolerable voice and data packet error probabilities, respectively. This multiple access capability limited by k shall be shared by all voice and data users. the CDMA channel is used to accommodate several voice calls simultaneously, while the data users follow the ALOHA protocol with retransmission control and contend for the remaining multiple-access capability of the channel. Several admission control scenarios are proposed. The key design issue is the amount of feedback global information about data backlog (N_t^d) and the number of established voice calls in progress (N_t^v) at time t;

$$\Phi(N_t^v, N_t^d) = (N_t^v, N_t^d).$$

A Markovian model of voice and data traffic is developed to evaluate the stationary distribution of $\Phi(N_t^v, N_t^d)$ which is used to evaluate the network performance measures such as call blocking, throughput, data packet delay, etc. However, this work assumes that the voice and data users have been already admitted in the system and the emphasis is on performance evaluation. There is no policy considered for admitting new users in the light of the system performance measures. In [74], the admission control problem was investigated in a similar approach besides the application of *linear* and *nonlinear* programming to ensure certain BER and maximum throughput. the interesting conclusion is that the error in channel measurement for actual number of active users has severe implication on the system performance.

Yang and Evaggelos [78] link system performance measures under currently carried users and the process of admitting new users into the network. An optimal admission control policy based on modeling the system operation by a semi-Markovian decision process (SMDP) is derived such that for the number of newly arrived voice users that are accepted in the network, the long-term blocking probability of voice calls is minimized. The results show that the optimal policy outperforms the direct one (users accepted as far as there are available CDMA codes) admission especially under heavy traffic load. Moreover, data traffic is modeled as an $M/D/c/K$ queue model. The new data users are rejected if the mean data delay (or the packet loss probability) exceeds a desirable prespecified level. Considering the other user interference of the CDMA network, two models were analyzed: the thresholds as described above and graceful degradation model where there is a nonzero probability of correct reception for any arbitrary number of packets (even for number of packets exceeding the threshold), which depends on the total number of simultaneously transmitted packets (data and voice). It was interestingly found that the data traffic performance depends drastically on the interference model used.

It is obvious that the above admission strategies do not provide a global solution for all potential users. In other words, they look at on individual basis such that the overall SIR at the base station input is not violated by accepting new user. Nevertheless, the power control problem is closely related to admission decisions.

Let a set of N users using a CDMA network which is composed of M cells. Then, the average system transmitted power is given by

$$I = \sum_{j=1}^{M} \sum_{i=1}^{n_j} S_{ij} \; G_{ij} \tag{4.9}$$

where G_{ij} is the gain loss associated with user i at base j, and S_{ij} is the average power of user i at base j. The objective of the power assignment, here, is to minimize I provided that certain QoS parameters are guaranteed. Therefore, the power assignment problem should be formulated as an optimization problem where the objective is to minimize the total transmitted power under certain constrains to find out whether it is feasible to accept a new user as shown below:

$$\text{minimize} \sum_{j} \sum_{i} S_{ij}$$
$$\text{subject to}$$

$$\begin{aligned} &\mathrm{SIR}_{ij} \geq \mathrm{SIR}_{ij}^{TH} && i = 1, \cdots, n_j, \; j = 1, \cdots, M \\ &S_i \geq 0 && i = 1, \cdots, n_j, \; j = 1, \cdots, M \end{aligned} \tag{4.10}$$

This problem integrated with base station assignment was investigated in [87] but the admission/congestion issue was not explicitly addressed. The work of Larijani and Hafez in [75] and [76] explicitly include the admission control and add another constraint that is

$$\mathrm{Prob}(\mathrm{SIR}_i \leq \mathrm{SIR}_{ij}^{TH}) \leq S_{ij}^{out}$$

where SIR_{ij}^{TH} is the signal-to-interference-ratio threshold for user i at base j and S_{ij}^{out} is the outage probability threshold for user i at base j. An optimum and non-optimum solutions were derived for the power allocation for each class of traffic such that each class attain the required QoS. Generally, these admission control schemes were found to be very effective on **Hot-cell** scenarios, where *intercell* interference is negligible compared to *intracell* interference.

4.4 Transmission Rate Control Based Admission policies

We shall start this section with the motivation behind the transmission rate based admission control. Following the approach used in [88], we can interrelate the desired SIR_i of a user i, transmission rate, R_i and total available bandwidth, W. Consider a single-cell system and ignore the background noise. From Eq. (4.1)

$$I = (n_i - 1)S_i + \sum_{j=1, j\neq i}^{N} n_j S_j \tag{4.11}$$

where n_i is the number of users operating at R_i, S_i is the user's power received at the BS side and N is the number of classes of users using the network. Assume $n_i \gg 1$, then each user will approximately suffer of the same amount of interference $I \approx \sum_{j=1}^{N} n_j S_j$. Now from Eq. (4.3), we can see the following relation:

$$\frac{S_1}{r_1 \mathrm{SIR}_1} = \frac{S_2}{r_2 \mathrm{SIR}_2} = \cdots = \frac{S_N}{r_N \mathrm{SIR}_N} = \frac{W}{\sum_{j=1}^{N} n_j S_j} \tag{4.12}$$

Using the relationship in Eq. (4.12), we get the following relation

$$\sum_{j=1}^{N} n_j R_j \mathrm{SIR}_j = W \tag{4.13}$$

Therefore, given that SIR_j is the QoS parameter required by user j, the desired performance of all users can be met if and only if Eq. (4.13) is satisfied. We can observe from Eq. (4.13) that there are two parameters which can be managed: 1) number of calls, n_j, 2) the transmission rate , R_j. Indeed, this is the philosophy behind this type of admission control schemes. Further, varying R_j may allow for more admission of users and hence the total capacity of the system has increased.

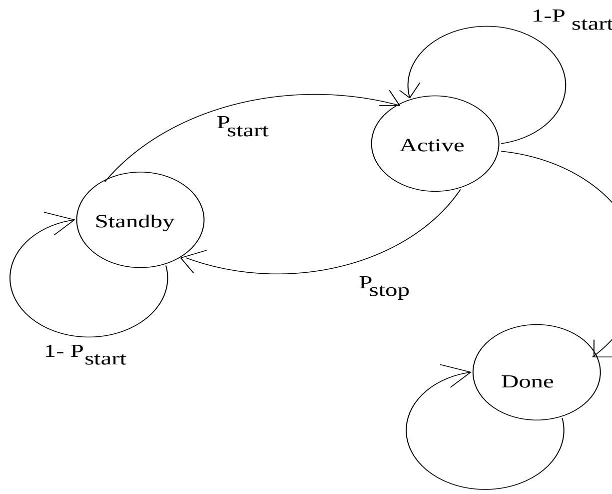


Figure 4.2: The state diagram of the packet level activity

The process of varying R_j is either "artificial" or actual change in user's information rate. The former type is based on modeling the user activities during **ON-state** period by three states: *active, standby and done.* During the active period, the user transmits information normally using R_j, while during the standby state the user stops transmitting information data and instead those information will be enqueued in the user's side. Therefore, the average user's transmission rate is less than R_j. So, the user's transmission rate has been adapted such that certain QoS requirements are met. Figure 4.2 shows the source model. By now, it should be clear that this class of admission policies depends on feedback information about the actual status of active users. Moreover, the bursty nature of the 3G wireless networks appeals for such type of call admission control.

The work in [72] assumes that certain voice and data calls are already admitted to the system and the proposed access scheme control the instantaneous flow of data packets into the network such that the outage probability condition is met most of the time. The outage probability is defined as the fraction of time the power assignment for active users is not feasible. The data packets flow is controlled by what is called *permission probability.* The permission probability is set by a non-negative integer parameter called the *persistence state.* If in the *nth* slot the persistence state, $T(n)$, is j, then each data user independent of other users, transmits with probability π^j and refrain from transmitting with probability $1 - \pi^j$. The parameter π $(0 < \pi < 1)$ is fixed probability and known to the data users. The persistence state is broadcast to all users by the end of each slot. It was shown that with appropriate choice of parameters, the persistence state scheme yield a simple, flexible and efficient method to tradeoff capacity and delay.

Liu and slivester [79] investigated a joint admission/congestion control for integrated CDMA network. The emphasis as usual is to guarantee certain quality (i.e. packet error rate (PER)) for real-time traffic while allowing narrow-time data traffic to utilize the residual channel capacity. The admission process of data traffic is composed of two phases. The first phase ensures the rejection of data calls which may experience long delay and it is achieved by not allowing more than a prespecified number of data users. The second phase ensures the PER of voice calls by limiting the transmission probability for data users. In other words, the active data user is modeled during its cell level by three states: standby, active, done as shown in Fig. 4.2.

In [83], an admission strategy is investigated where a joint call level/packet level admission control is proposed. The important part of this strategy is where a fixed/adaptive scheduling is imposed by the base station to minimize the intracell interference and the same time try to maximize the distance among the users with higher energy per bit. At call level, a modified equivalent bandwidth technique is employed. By this algorithm, each user call is assigned an index and request for connection is granted based on the sum of these indices. As a matter of fact, this set of indices is a function of interference variance which will have lower value due to the employment of delay scheduling. The proposed strategy was tested under three types of traffic (voice, fixed rate video and variable rate video) and proved to be effective in providing lower call blocking probability. In this work, it is assumed that the users activities are well known ahead or the estimation of these parameters is perfect.

In [84], a simple admission algorithm where the power and data rate of mobile data users are adjusted to allow more users on a congested network is investigated. The key aspect of this algorithm is to impose delay on data users such that the call blocking is minimized. For a user (voice or data) attempting to access the system when the equivalent number of full rate channel is m (maximum number of servers), all data users will lower their power levels and data rates proportionally. Defining $R_n \geq R_{min}$, where R_n and R_{min} are the normalized and minimum normalized data users transmission rate, respectively. R_n is a function of the number of users in the network as follows, i.e.

$$R_n = \begin{cases} \frac{m-k_v}{k_d} & , \frac{m-k_v}{k_d} \geq R_{min} \\ R_{min} & , \text{otherwise} \end{cases} \tag{4.14}$$

where k_v, k_d are the number of active voice and data users respectively. Voice users are always assumed to work at full rate. If all data users are already at the lowest allowable data rate, R_{min}, the call is blocked. This algorithm was studied under assumption of perfect power control [84] and imperfect power control [85]; all modeled the system as an $M/M/m$ Erlang loss system and all data users operate at the same rate at time t. The results show significant improvement in call blocking rate and even lower waiting delay under low traffic load. Moreover, this algorithm is proved to be sensitive to the imperfections in the power control.

In [89], the voice interference $Z(t)$ is measured periodically and data users continuously vary their transmission rates accordingly so that more data users can be accommodated. The fact is that it is very difficult to implement such scheme. A modified version has also been proposed where a fixed rate is assumed and the number of data users allowed to transmit is controlled by the level of voice interference (i.e. $k_d = \Psi(Z(t))$. k_d is determined periodically. If the number of in-service data users (i.e k_d) in the previous measurement interval is greater than k_d for the current measurement interval, the oldest in-service data users are notified to stop transmission and should join a sub-queue.

4.5 Combined Power / Rate Control Based Admission policies

As we have mentioned above, the two entities (*power and rate*) are very related to each other. Some of the techniques in section 4.4 have assumed that power control is applied. However, they do not include this crucial element of CDMA system in the call admission policy. Therefore, it is expected that including both elements (*power and transmission rate*) in the admission policy should improve the system performance. Furthermore, simulation results show that call dropping decreases if maximum power constraint on a BS basis instead of on a per channel basis [82]. Therefore, the problem of resources management could be formulated as an optimization problem to maximize the total throughput of the network [80]-[82].

Here is the problem formulation:

$$\underset{p}{\text{maximize}} \quad \sum_{i=1}^{N} R_i$$

subject to

$$\begin{aligned} & S_{tot,0} \leq S_{totMax,0} & & \\ & R_i \geq R_i^{min} & & i = 1, \cdots, N \\ & \mathrm{SIR}_i \geq \mathrm{SIR}_i^{TH} & & i = 1, \cdots, N \end{aligned} \tag{4.15}$$

All above mentioned work focus on the uplink channel. Though, non-real time services such as WWW-browsing sets high requirements on the capacity, especially on the downlink. However, this admission algorithm is fixed in the sense that it does not take into consideration the activity of admitted users. More, the complexity inherited in this algorithm may be justified for the superior throughput performance.

Combined power/rate control was also proposed in [85] where only data users are assumed in the system and data users reduce their transmission rate continuously such that more data users can be accommodated. It was shown that about 231% increase in the total transmission rate was achieved, when the multipath fading consists of a single Rayleigh path. Again, the difficulty issue of implementing such scheme raises here.

4.6 Effective Bandwidth Based Admission Control

Referring to Eq. 4.13, we can put this equation in a more generalized form [90], i.e

$$\sum_{j=1}^{N} n_j e_j \leq W \tag{4.16}$$

where $e_j = R_j \mathrm{SIR}_j$ is called the *effective-bandwidth* factor. For example, let R_1=10 Kbps and the required SIR_1=7 dB, then e_1=50 kHz which means that the effective bandwidth used or should be reserved for such user is 50 kHz out of the total available bandwidth. We can observe from Eq. (4.13) that there are two parameters to play with: 1) number of calls, n_j, 2) the transmission rate , R_j. Indeed, this the philosophy behind this type of admission control scheme.

4.7 Conclusions

This chapter has discussed briefly the existing admission/congestion techniques for CDMA networks. The philosophy behind most of these techniques is to control either the user's transmission power, transmission rate, or both such that the required QoS (i.e. BER only) is maintained. Nevertheless, these power control techniques will not be sufficient taking in consideration the characteristics of the future network users and their QoS requirements such as (BER, delay, packet blocking, etc.). Hence, the design of any future admission/congestion policy should consider these performance measures. Moreover, it is important for any future admission/congestion policy to combine between the preventive and reactive mechanisms.

Chapter 5

Hybrid MC-CDMA/TDMA Utilizing Decorrelator Receiver

5.1 Introduction

The essence of this work is to introduce an interaction between the physical layer and higher layers, enabling a more practical utilization of multiuser detection and supporting services with different QoS parameters. To achieve this objective, two traffic flow control approaches accompanied a hybrid TDMA/MC-CDMA system utilizing multiuser detection are proposed here. One approach deterministically controls the flow of traffic into the TDMA slots [93], while the other statistically controls the flow of traffic depending on the instantaneous changes in the traffic load. Although, hybrid TDMA/CDMA techniques were proposed in the literature before (see e.g. [12]), to our knowledge, our proposal is new and original at least by creating an interaction between the physical layer and other higher layers (for example, ATM layer and or the flow and congestion control functions in the data links or network layers) that could enable the integration of multimedia applications into wireless networks.

The rest of the chapter is organized as follows. The proposed multiple access protocol is presented in section 5.2. Then, section 5.3 looks at the statistical modeling of the traffic sources considered in this work. Then, the system performance analysis is presented in section 5.4. In section 5.5, we present and discuss the results and the performance of the proposed system. Conclusions are drawn in section 5.6.

5.2 Proposed Multiple Access Protocol

The interaction between the physical layer and higher layers is of great interest for all third generation communication systems which support services with different QoS parameters. In this work, the hybrid TDMA/CDMA is adopted as the physical layer where the time domain is divided into frames and each frame is composed of just two time slots (i.e. T_{s1}, T_{s2}). However, the detection strategy for the population of each time slot is different. The flow of traffic to each time slot is controlled by the

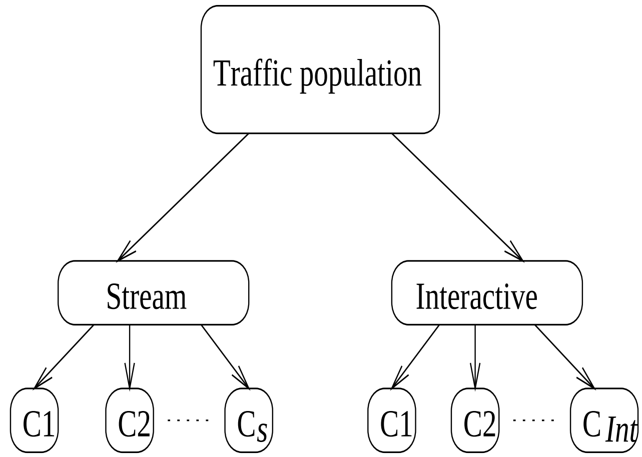


Figure 5.1: Classification of traffic population

traffic characteristics of each user, its QoS parameters and the detection strategy applied in that time slot. Therefore, we propose the following classification for the traffic population. The traffic is divided into two categories: **Stream Traffic and Interactive Traffic**. Each category has classes of traffic which have common traffic characteristics. The stream traffic includes classes of traffic that have high transmission activities. Nevertheless, it includes different types of users where each one has its own transmission characteristics (i.e average rate, peak rate, activity of the transmission, etc.). On the other hand, the interactive category of traffic comprises all other traffic that are not included in the first category (e.g. signaling, high bursty users) as well as the excess traffic of the stream traffic (i.e. Variable bit rate components) which will be explained in the next subsection. Figure 5.14 shows these classifications, where stream category has C_s different classes, while the interactive traffic category has C_{Int} different classes.

The justification for this classification relies on the fact that each category has its own required QoS as well as its own traffic and transmission characteristics. Hence, each category of traffic should be treated differently. Next, we present the flow traffic control scheme from both traffic categories.

5.2.1 Traffic flow control

It is well known that the performance of CDMA techniques (in particular DS-CDMA) is very sensitive to the processing gain. Hence, the user's transmission rate is not expected to be high enough to support multimedia applications nor a wireless ATM system where the objective is on-demand availability of bandwidth at peak rate as high as 10 Mb/s [1]. As an example, let the available bandwidth be 20 MHz (which is the maximum proposed bandwidth for future **Wideband CDMA** system [6]), and the processing gain 100, then the maximum bit rate that can be supported is 0.2 Mb/s. On the other hand, the maximum bit rate is more than 0.2 Mb/s using multi-code techniques. This is the force behind our proposal for high rate users.

Considering the heterogeneous traffic of the future communication networks, and for analysis convenience, we propose that various traffic classes should represent their bit rates as n multiple of a basic rate R_b; or $1/n$ multiple of R_b; n is an integer. For practical rates, users of rate R_b/n, the generated packets will be buffered till their gross rate matches with R_b. To utilize the bandwidth at our disposal efficiently, and at the same time invest the capability of the CDMA technique, it is suggested that each class of users has its own basic rate which is chosen according to the QoS requirements. For example, a class of users that can tolerate a 10^{-3} bit error rate might not be effective for their transmission rates to be subdivided by the same reference rate (i.e. basic rate) of another class of traffic that its bit error rate requirement is much smaller than 10^{-3}. For high rate classes (say class j, where the bit rate is R_j), when a user needs to transmit with a rate greater than R_b, the high-bit stream shall be converted into n low-bit basic streams, where $n = R_j/R_b$. Each new low-bit basic stream has a bit rate equal to R_b. Consequently, the low-bit basic stream is packetized into packets of fixed size. This strategy of subdividing the high rate users into low-rate streams and then applying the CDMA technique is called Multi-Code CDMA (MC-CDMA). MC-CDMA has very interesting features that the

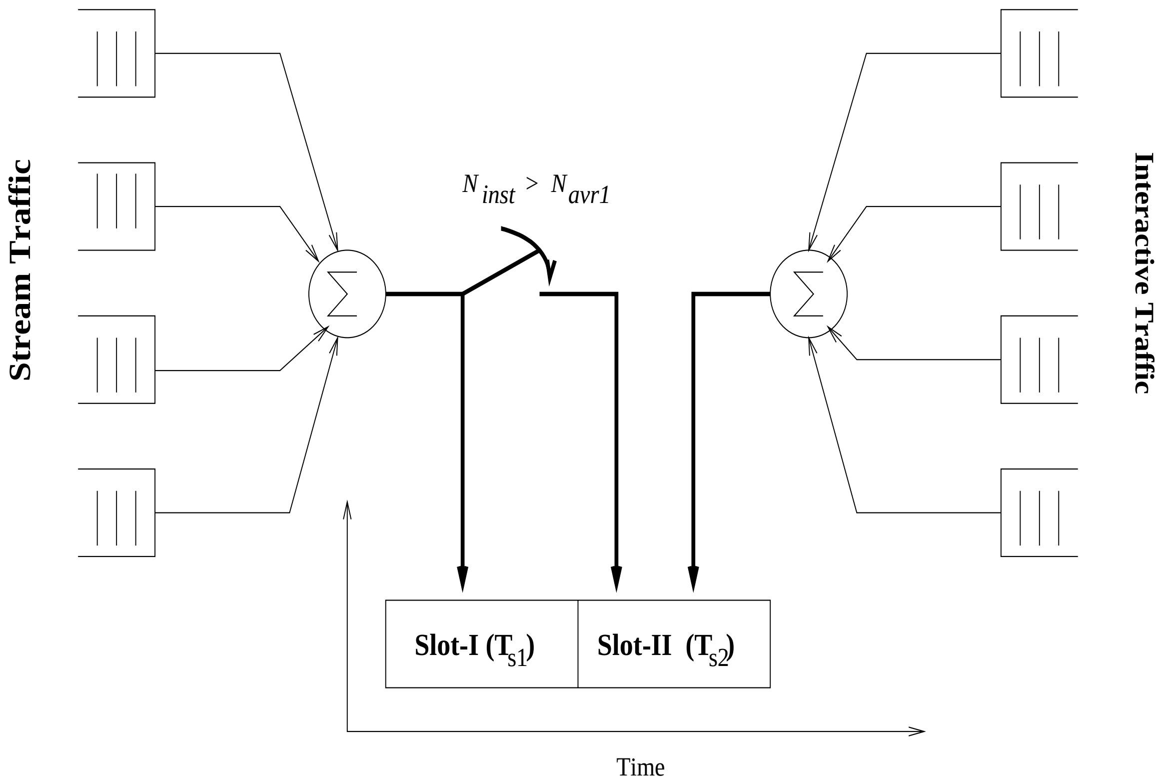


Figure 5.2: The proposed multiple access protocol.

other systems lack. It can easily integrate traffic of different transmission rates in a unified structure where all the transmissions over the radio channel occupying the same bandwidth and having the same processing gain [91].

This unified approach can easily be generated in practice using symbol repetition with or without puncturing and sequence repetition. As a matter of fact, a similar notion was suggested for CDMA2000 to support variable bit rate users [92].

The MC-CDMA technique is general and it can be implemented on top of any switching protocols such as circuit switching, packet switching, ATM, etc. For instance, consider the ATM technology, the accommodation of VBR transmission through implementation of MC-CDMA along with the proposed traffic control policies guarantees relative simplicity and speed compared to ATM accommodation mechanism of VBR users. However, there is nothing in our new technique that could forbid the added benefits of ATM VBR technique on top where the proposed techniques in this paper could be adopted in the ATM adaptation layer (AAL).

The interesting implication of using MC-CDMA is the huge increase in the power level of the total interference. Hence, the QoS will degrade severely as the number of active users increase. To mitigate these sever degradation in performance, we propose the use of hybrid two slots TDMA/MC-CDMA where the detection strategy is different for each time slot population. In addition, the number of simultaneous packets emitted from interactive traffic users is expected to be low due to the high burstiness of such kind of users. Therefore, the *Conventional Receiver* (i.e single-user DS-CDMA receiver) is suitable and shall be employed for the demodulation of signals received during T_{s2}. Further, since the *Decorrelator Multiuser Receiver* has very good theoretical performance irrespect of the number of instantaneous users, we shall employ it for the users of T_{s1} (i.e. stream traffic) where the users activity is high. However, as it was mentioned earlier that the lack of knowledge about the actual number of packets on the channel will also severely degrade the system performance. Hence, the main objective of our traffic flow control approaches is to make the implementation of the hybrid TDMA/MC-CDMA utilizing multiuser detection possible and more practical by forcing the number of packets on T_{s1} to be completely known to the receiver.

The 1st Approach [94]

Basically, T_{s1} is dedicated to the stream traffic users, while, T_{s2} is dedicated to the interactive traffic users. During the signaling period (where the user tries to communicate with the base station to get an admission to the system), the stream traffic user j negotiates with the base station about (among other things) the average transmission bit rate $R^s_{avr(j)}$ that the user shall stick to while it is using T_{s1}, though the actual bit rate might be less in some cases. In the case that the instantaneous rate is less than $R^s_{avr(j)}$, the user has to stick to the agreement and generate dummy packets. Of course, these dummy packets will affect the bandwidth utilization.

However, since this category of users has high transmission activity, it is expected that its role will me minimal.

Now, it is clear that this fixed rate is translated into a fixed number of received packets (i.e. ${N^s}_{avr}$ as in Eq. 5.1) at the receiver side where a multiuser detector is used. Thus, this simple traffic flow control enables a practical and effective implementation of Multi-User Detection strategy at the receiver of the stream traffic population since the numbers and identities of overlapping CDMA packets are perfectly known through out the user call.

$$ {N^s}_{avr} = \sum_{j=1}^{N^s} {N^s}_{avr(j)} = \sum_{j=1}^{N^s} {\hat{\xi}^s}_j {n^s}_j \tag{5.1} $$

where ${N^s}_{avr}$ is the average number of packets emitted from all active stream users to T_{s1}, N^s is the number of classes of users in the stream traffic category, n_j^s is the number of active stream users belonging to class j, and ${\hat{\xi}^s}_j = R^s_{avr(j)}/R_b$ is the ratio of a user's average bit-rate to the basic bit-rate. Now, if the user needs to send information using a bit rate higher than the average bit rate (i.e. $R_i^s > R^s_{avr(j)}$), then the excess packets (variable bit rate components) should be queued and directed to the other time slot (i.e. T_{s2}) where these signals along with interactive traffic flow shall be detected by the conventional receiver (i.e. single user). Therefore, the assignment of slots to a part or all of user traffic is done apriori through an agreement between the base station and the user.

$$ {N^s}_{excess} = {N^s}_{inst} - {N^s}_{avr} \tag{5.2} $$

where ${N^s}_{inst}$ is the instantaneous number of packets emitted from active stream users.

The 2nd Approach [95]

It is obvious that the 1st approach heavily depends on the availability of call signaling (i.e. valid only for connection oriented calls). The other alternative is to control the flow of traffic according to the instantaneous changes in the traffic population. Again, T_{s1} is dedicated to the stream traffic users, while, T_{s2} is dedicated to the interactive traffic users. The difference between the two approaches manifests itself in the following notion. The base station (BS) is constantly monitoring the traffic load initiated from both traffic categories. At the moment the stream traffic load exceeds the average traffic load as defined in Eq. (5.3), the BS must send signals to the new users belonging to the stream traffic category who desire to transmit that they should use the other time slot (i.e. T_{s2}).

$$N^s{}_{avr} = \sum_{j=1}^{N^s} N^s{}_{avr(j)} = \sum_{j=1}^{N^s} \xi^s{}_j n^s{}_j \theta^s_{j,on} \tag{5.3}$$

Where $\theta^s_{j,on}$ is the activity factor of a user from class j; a measure of burstiness to be defined shortly, $\xi^s{}_j = R_j/R_b$ is the ratio of a user's bit-rate to the basic bit-rate. On the other hand, dummy packets will be transmitted by users on command of BS at T_{s1} if user traffic is less than $N^s{}_{avr}$ as defined in Eq. (5.3). Those bulk of users directed to T_{s2} is called **Excess Traffic** which is derived as in Eq. (5.4).

$$N^s{}_{excess} = N^s{}_{inst} - N^s{}_{avr} \tag{5.4}$$

It should be noted that if Eq. (5.2) or Eq. (5.4) results a negative value, it means that there is no excess traffic.

Figure 5.2 illustrates the proposed communication protocol in the TDMA/MC-CDMA transmission scheme. At this point, it is clearly seen that the first approach resembles deterministic traffic control, which costs more in network capacity than the second approach, where we try to statistically but perfectly estimate the instantaneous amount of packets on the channel. In fact, the second approach is amenable to packet connectionless mode of operation (i.e. no call establishment phase).

Having introduced these traffic flow control approaches that enables the practical implementation of the *Decorrelator Multiuser* receiver, we should note the following. The overall average performance will never be better than a pure theoretical decorrelator receiver performance, but always better than the conventional (single user) performance. It is the best compromise that will make the implementation possible.

However, a significant limitation of this technique is the computational complexity due to the inversion of the correlation matrix [45] where entries depend on the number of active users, signature sequences, and the delays of the users. Further, any change in one of these parameters changes the correlation matrix and consequently a need for updating the multiuser detection process. Moreover, the uncertainty in the actual number of active users is another serious problem that might degrade the system performance very severely [66] which now could be resolved by applying our 1st approach for traffic control. Considering the 2nd approach for traffic control, we observe that the accuracy of the knowledge of actual number of transmitted packets is dependent on the estimated statistical average. In addition, this estimated statistical average might change during the user call and, consequently, change in the correlation matrix. However, these changes in the correlation matrix will be much less if it is compared to pure CDMA where the correlation matrix will be vulnerable to a wide range of statistical changes of traffic sources (e.g. transitions from ON state to OFF state and vice versa).

5.3 Statistical Traffic Models

The structure of the voice and video traffic is fairly complex due to the high correlation among arrivals [98]. Furthermore, the activities of such sources play a key role in modeling generated packets. The correlation between voice packets generated during a call can be modeled as an Interrupted Poisson Process (IPP) (also it is call ON/OFF model) as shown in Fig. 5.3. This model is the simplest model and is widely used to model voice traffic. This model allows relevant parameters such as maximum packet rate of source, mean packet rate and mean duration to vary

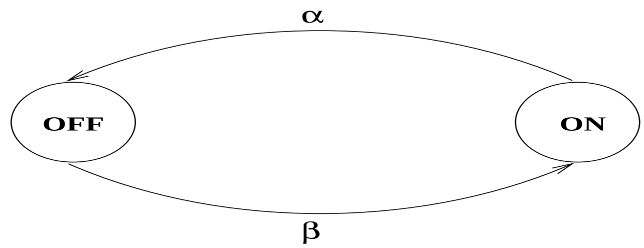


Figure 5.3: MMPP traffic source model.

independently of each other. In fact, the Interrupted Poisson Process (IPP) is a special case of a Markov Modulated Poisson Process (MMPP) where no packets are generated at the idle state while during active bursts the source transmits packets at its peak rate σ_v. While this model does not take into consideration the correlation in voice or data bursts, sources with such correlation can be shown to be resolved into a number of such mini sources as that in Fig. 5.3, and thus no loss of generality is encountered in assuming a multitude of mini sources such as that of Fig. 5.3 as representative of the activities of the accepted calls.

Furthermore, recent results [97] based on experimental measures have shown this model to be also more appropriate for data packets generation than the Poisson approximation which does not capture any correlation between consecutive packets arrival. Therefore, in this work, the ON/OFF model will be also used for modeling all traffic users during the active call period.

The activity of each call follows this state diagram. β is the rate transition out of OFF state. While α is the rate transition out of ON state. The active and idle periods are assumed to be Geometrically distributed with the parameters above. Solving the state diagram of Fig. 5.3 (writing the balance equations), one gets the probability of a call from k user being in active burst (transmitting packets), or silence state (no packets transmitted), respectively; i.e.,

$$\theta_{k,on} = \frac{\beta_k}{(\alpha_k+\beta_k)}, \quad \theta_{k,off} = \frac{\alpha_k}{(\alpha_k+\beta_k)} \tag{5.5}$$

The total number of packets generated from all active calls (during the ON period) follows a Bernoulli distribution.

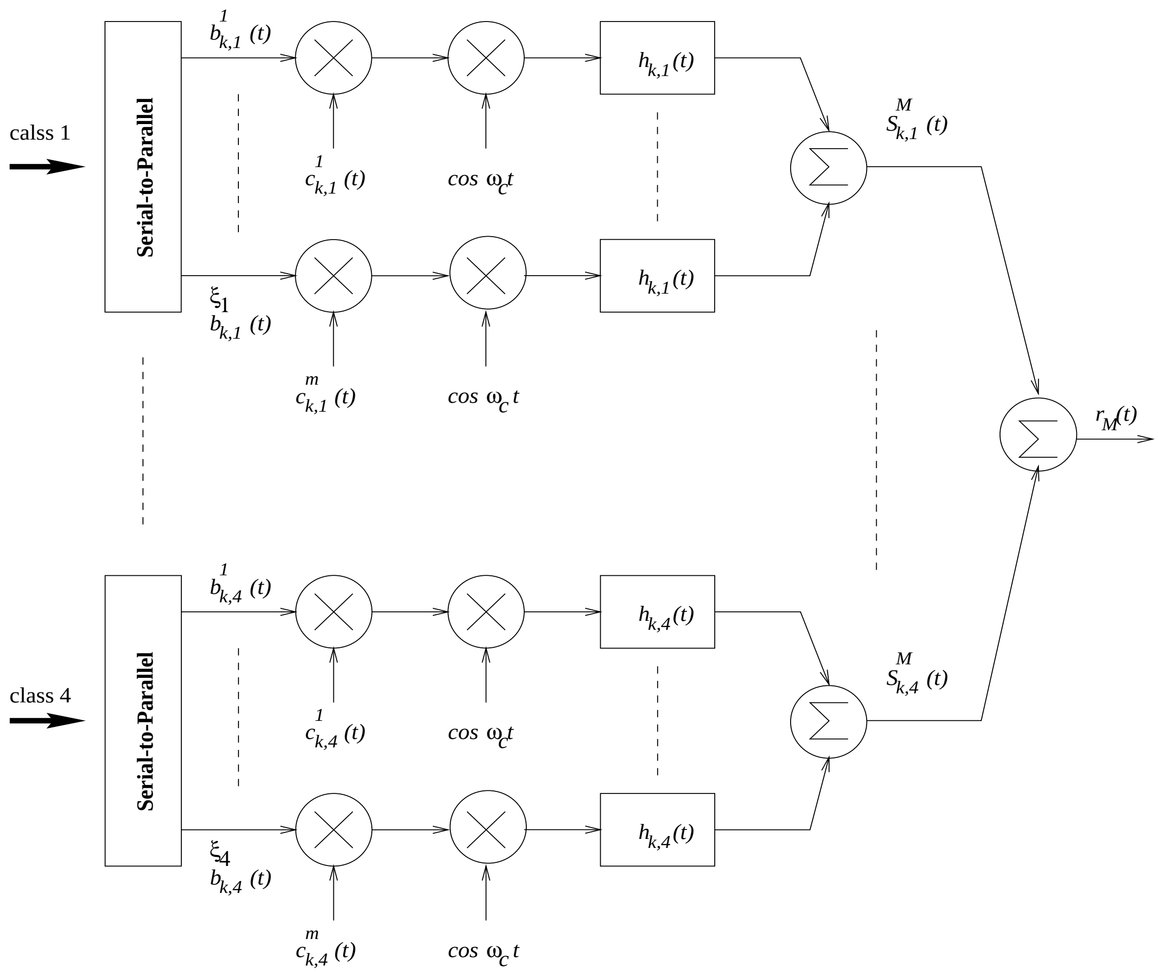


Figure 5.4: The general structure for MC-CDMA transmitter.

Considering that I different classes of users are multiplexed, the aggregated packet arrivals are governed by the number of calls in the ON state from all types of active users. Having assumed that all classes of traffic users are independent identically distributed random variables (iid), the compound traffic distribution of i packets generated by all active calls (i.e. λ_i, $i = \sum_{k=1}^{I} n_k$) of all I classes is given by

$$\lambda_i = \sum_{n_1=0}^{N_1} \sum_{n_2=0}^{N_2} \cdots \sum_{n_I=0}^{N_I} \prod_{k=1}^{I} P_k(n_k) \tag{5.6}$$

where N_k is the maximum possible number of packets generated by class k and $P_k(n_k)$ is the probability distribution of generating n_k packets from class k which is given by

$$P_k(n_k) = \begin{pmatrix} N_k \\ n_k \end{pmatrix} (\theta_{k,on})^{n_k} (1 - \theta_{k,on})^{N_k - n_k} \tag{5.7}$$

5.4 System Performance Analysis

In the following, we analyze the performance of the proposed TDMA/MC-CDMA system employing the above traffic control approaches. The analysis applies equally for both categories of traffic unless it is stated differently. Hence, and for analysis convenience, the system parameters are defined generally. Figure 5.4 shows the general structure of multicode CDMA transmitter. The admitted users into the communication network will comply with the agreement with the BS. If a user need to transmit its information using higher transmission rate R_j than the basic rate R_b, the *serial-to-parallel* unit shall split the coming information stream into $\xi_j = R_j/R_b$ low-bit basic streams. Each new low-bit basic rate has a bit rate equal to $1/T_b$, where T_b is the basic bit duration. $T_b = \xi_j T_j$, where T_j is the bit duration of the high-bit rate stream traffic coming from class j. For example, assume there are four users, $R_1 = 10$ Kbps, $R_2 = 10$ Kbps, $R_3 = 20$ Kbps, and $R_3 = 40$ Kbps, where all of them belong to the same class of traffic. Let $R_b = 10$ Kbps. Thus, packets coming from user1 and user2 will be transmitted as they are, while the packets coming from user3 and user4 are going to be subdivided (multiplexed) into 2 and 4 new packets, respectively. Then each low-bit rate $b^i_{k,j}(t)$ will be spread with a different gold code $c^i_{k,j}(t)$, where $c^i_{k,j}(t)$ is the spreading signal associated with the ith low-bit rate stream of the kth user belonging to the j class of traffic. In general, the spreading code consists of a sequence of positive and negative pulses; $d_l \in \{-1, 1\}$.

$$c(t) = \sum_l d_l \Psi(t - lT_c) \tag{5.8}$$

where Ψ is a rectangular waveform of duration T_c. Similarly, $b^i_{k,j}(t)$ is the baseband data signal that consists of a sequence of positive and negative pulses; $a_m \in \{-1, 1\}$.

We define the information bits stream generally as

$$b(t) = \sum_{m} a_m \Phi(t - mT_b) \tag{5.9}$$

where Φ is rectangular waveform of duration T_b. Usually $c(t)$ has a much larger bandwidth than $b(t)$. Then, all low-bit rate BPSK baseband signals (i.e. $b^i_{k,j}(t) \oplus c^i_{k,j}(t)$) belonging to the kth user will be modulated with the carrier frequency ω_c and transmitted simultaneously. Therefore, all ξ_j low-bit streams belonging to the kth user are experiencing the same channel $h_{k,j}(t)$, where $h_{k,j}(t)$ is the channel impulse response.

Now, it is easy to represent user's signals admitted into the system as follows,

$$S^M_{k,j}(t) \;=\; \sum_{i=1}^{\xi_j} \Big(c^i_{k,j}(t) b^i_{k,j}(t) cos(\omega_c(t) + \phi_k) \Big) \otimes h_{k,j}(t) \tag{5.10}$$

where $S^M_{k,j}(t)$ is the compound signals coming from the kth user belonging to the j calls using the Mth slot for transmission; $M = 1, 2$. Further, ξ_j is R_j/R_b, and not to mention that ξ_j is limited by the practical application (for example, the existing spreading codes, etc.), ϕ_k is the phase shift associated with the kth user.

Nevertheless, under each approach and for each time slot, the actual number of low-bit basic rate streams ξ_j will be different. First, consider the received signals in T_{s1} (i.e. $r_1(t)$) which is composed of a part of the stream traffic as explained before. Assume that there are four classes of stream traffic. Hence,

$$r_1(t) = \sum_{j=1}^{4} \sum_{k=1}^{k^s_j} S^1_{k,j} + n(t) \tag{5.11}$$

where $\xi_j = \hat{\xi^s}_j = R^s_{avr(j)}/R_b$ (i.e. $R^s_{avr(j)}$ is either the agreed upon average rate as in the 1st approach or the statistical average rate as in the 2nd approach). $n(t)$ is an additive white Gaussian noise with two sided spectral density $\frac{N_o}{2}$ W/Hz. k^s_j is the number of active stream users belonging to class j of the stream traffic category and it is a random variable changing according to the number of accepted calls from the j class. On the other hand, the received signals in T_{s2} (assuming there are two interactive classes of users) are composed of two parts: 1) the interactive users signals and 2) the excess traffic from the stream users. Thus,

$$r_2(t) = \sum_{j=1}^{2} \sum_{k=1}^{k_j^{Int}} S_{k,j}^2(t) + \sum_{j=1}^{4} \sum_{k=1}^{k_j^s} S_{k,j}^1(t) + n(t) \tag{5.12}$$

where k_j^{Int} is the number of active interactive users belonging to class j of the interactive traffic category and it is a random variable changing according to the number of accepted calls from the j class. Now, regarding $S_{k,j}^2(t)$, $\xi_j = \xi_j^{Int} = R_j^{Int}/R_b$. However, if we consider the second summation, we recognize that the variable components (i.e excess traffic) of high rate stream users will be split according to Eq. (5.13).

$$\xi_j = \frac{R_j^s - R_{avr(j)}^s}{R_b} \tag{5.13}$$

Having transmitted the MC-BPSK-CDMA signals, they will be detected according to the proposed approaches. In other words, $r_1(t)$ is processed by the decorrelator multiuser receiver, while $r_2(t)$ is processed by the conventional detection strategy (single-user receiver).

In the following, we shall evaluate the packet error performance of the proposed approaches assuming there is no fading and the transmission of packets is **synchronous**. Following the above traffic control approaches, the stream traffic packets experience two different transmission strategies. Therefore, the bit error performance is composed of two portions. Firstly, the bit error performance P_{e1}^s due to packets using T_{s1} and detected by the multiuser receiver.

$$P_{e1}^s = P_e(N_{avr}^s, 1) \tag{5.14}$$

where $P_e(N_{avr}^s, 1)$ is the bit error probability when there are N_{avr}^s overlapping packets using T_{s1} and is given by Eq. (3.20). Secondly, the bit error performance $P_{e2}^s(i)$ due to the existence of i excess traffic packets using T_{s2} and detected along with the interactive packets by the single-user receiver.

$$P_{e2}^{s}(i) = \sum_{l=0}^{sig} \lambda_l^{Int} P_e(l+i,2) \tag{5.15}$$

where sig is the instantaneous number of interactive packets. This summation weighs the influence of the coexisting interactive packets on the excess of stream traffic by λ_l^{Int} as defined in Eq. (5.6), and $P_e(l+i,2)$ is the bit error performance when there are $l+i$ overlapping packet using T_{s2} and is given [53] by Eq. (5.16).

$$P_e(k,2) = Q\left(\sqrt{(E_k/N_o)^{-1} + \frac{k-1}{3PG}}\right) \tag{5.16}$$

where PG is processing gain (i.e. $PG \approx T_b/T_c$) and we define $k = l+i$ for convenience. Now, if there is no excess traffic (i.e. $N_{inst}^{s} \leq N_{avr}^{s}$), the bit error performance is just due to P_{e1}^{s}. Otherwise, the errors occur in the transmitted packets are due to P_{e1}^{s} as well as $P_{e2}^{s}(i)$. Hence, the correct bit-decision $P_c^{excess}(i)$ in this case is given by

$$P_c^{excess}(i) = (1 - P_{e1}^{s})(1 - P_{e2}^{s}(i)) \tag{5.17}$$

Then, this correct bit-decision should be weighted by the likelihood of the occurrence of i excess traffic packets (i.e. $\lambda_{j+N_{avr}^{s}}^{s}$ as defined in Eq. (7.8)). Consequently, the average bit error probability for the whole stream traffic category can be derived as follows.

$$P_e^{s} \;=\; P_e(N_{avr}^{s},1)\Lambda_{avr}^{s} + \sum_{j=1}^{N_{excess}^{s}} \lambda_{i+N_{avr}^{s}}^{s}(1 - P_c^{excess}(i)) \tag{5.18}$$

Though this discussion is applied for both approaches, each approach has different overall bit error probability. This manifests its self in Λ_{avr}^{s} which is a modified stream traffic distribution due to the proposed traffic control. Considering the first approach, It is always assured that there are N_{avr}^{s} packets transmitted in T_{s1}, so $\Lambda_{avr}^{s} = 1$. On the other hand, in the second approach, the coexistence of N_{avr}^{s} overlapping packets is dependent on the estimated statistical average. Consequently,

Λ^s_{avr} is defined as Eq. (5.19)

$$\Lambda^s_{avr} = \begin{cases} 1 & ; N^s_{inst} \leq N^s_{avr} \\ \lambda_{N^s_{avr}} & ; \text{otherwise} \end{cases} \tag{5.19}$$

Hence, it is obvious that for users whose transmission rates are less than or equal to the estimated average transmission rates, the probability is unity, because the user should comply with transmitting N^s_{avr} packets whether it actually needs to transmit that mount of packets or not. It is clear from Eq. (5.18) that the stream traffic bit error performance will be a compromise between two performance bounds (i.e multiuser receiver performance and single-user receiver performance). It will never be better than the first, but always better than the later (single-user receiver).

Following the same argument above, we obtain a similar equation for the average bit error performance of the interactive traffic users; P^{Int}_e.

$$P^{Int}_e = \sum_{j=1}^{sig} \lambda^{Int}_j \sum_{l=0}^{excess} \lambda^s_{l+N^s_{avr}} P_e(l+j,2) \tag{5.20}$$

We compare our proposal to the conventional system, where both categories of traffic are using the whole available bandwidth and the received signals are processed by the conventional single-user receiver. Further, both systems (proposed and conventional) are compared (when applicable) to the single user system (where there is just one user) which is considered the lower bound for any Multiple Access System. The common criterion used to evaluate the performance of the multiple access technique is the packet error rate, where each packet is composed of 53 bytes, and the packet error rate is then given by,

$$P_{packet} = 1 - (1 - P_e)^n \tag{5.21}$$

where P_e is the bit error probability (i.e. P^s_e for stream traffic users or P^s_e for interactive users), and n is the number of bits in the packet.

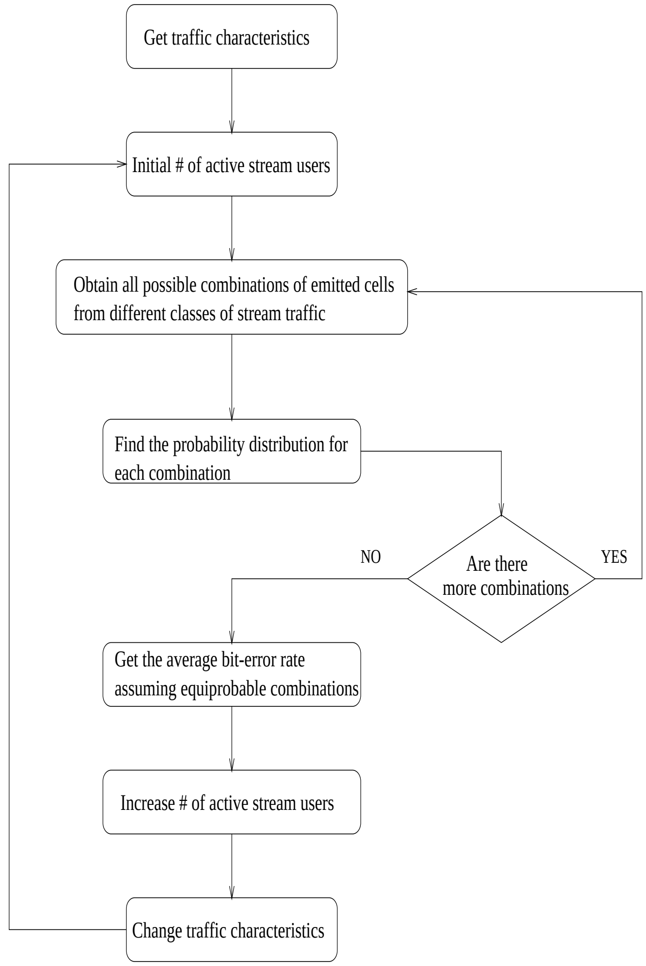


Figure 5.5: The flow chart for the error rate calculation.

5.5 Results

The results presented here are of the MC-BPSK-CDMA/TDMA hybrid system. Figure 6 illustrates how the packet error rate is evaluated. First, for every set of stream traffic parameters, and for every fixed number of active users, we obtain all possible combinations of packets that each class of users might send and each user will split its stream of bits into certain number of low-rate streams. It is worth emphasizing that the number of active users admitted into the system is not necessarily equal to the number of mutual packets processed by the system, because the number of

Table 5.1: Traffic Characteristics and System Parameters.

	Stream Traffic # of Classes= 4	Interactive Traffic # of Classes= 2
Class 1	R_1^s= 10-30 kbps, $R_{avr(1)}^s$=20 kbps # of users= 50, $\theta_{1,on}^s$= 1.0	R_1^{Int}=100 kbps # of users= 3,$\theta_{1,on}^{Int}$= 0.038
Class 2	R_2^s= 130-500 kbps, $R_{avr(2)}^s$= 260 kbps # of users= 3,$\theta_{2,on}^s$= 0.75	R_2^{Int}= 40 kbps # of users= 10,$\theta_{1,on}^{Int}$= 0.35
Class 3	R_3^s= 60-380 kbps, $R_{avr(3)}^s$= 160 kbps # of users = 5, $\theta_{3,on}^s$ =0.90	
Class 4	R_4^s= 100-640 kbps, $R_{avr(4)}^s$ = 290 kbps # of users= 2 ,$\theta_{4,on}^s$= 0.80	
Bandwidth, BW	10 MHz	10 MHz
Basic rate, R_b	10 kbps	10 kbps
Spreading code	gold code	gold code
Processing gain, PG	512	512
Packet size	53 bytes	53 bytes
Maximum # of users	60	13

packets depends on other parameters and not just the number of active users. Furthermore, each low-rate stream will be spread using a set of pseudo-random codes (Gold codes) where PG=1023. Though, in the case of two slots CDMA/TDMA and because of the 2 slots framing, the bursty rate should be higher than $1/T_b$. Here, we assume that the bursty rate at each time slot is $2/T_b$. Hence, to maintain the same bandwidth, the processing gain should be half the one used in the wide CDMA (i.e. conventional system in our case), that is 512. This way, the bit energy for each low-bit rate E_b is the same for both systems (i.e. conventional and proposed). The bit error rate is calculated using the classical multiuser detection and single user detection and averaged over all the possible combinations assuming that each combination is equiprobable. This is very important, because each class of traffic has different traffic characteristics. Hence, to obtain a fair and clear picture of the system performance, the results should be averaged over all possible combinations.

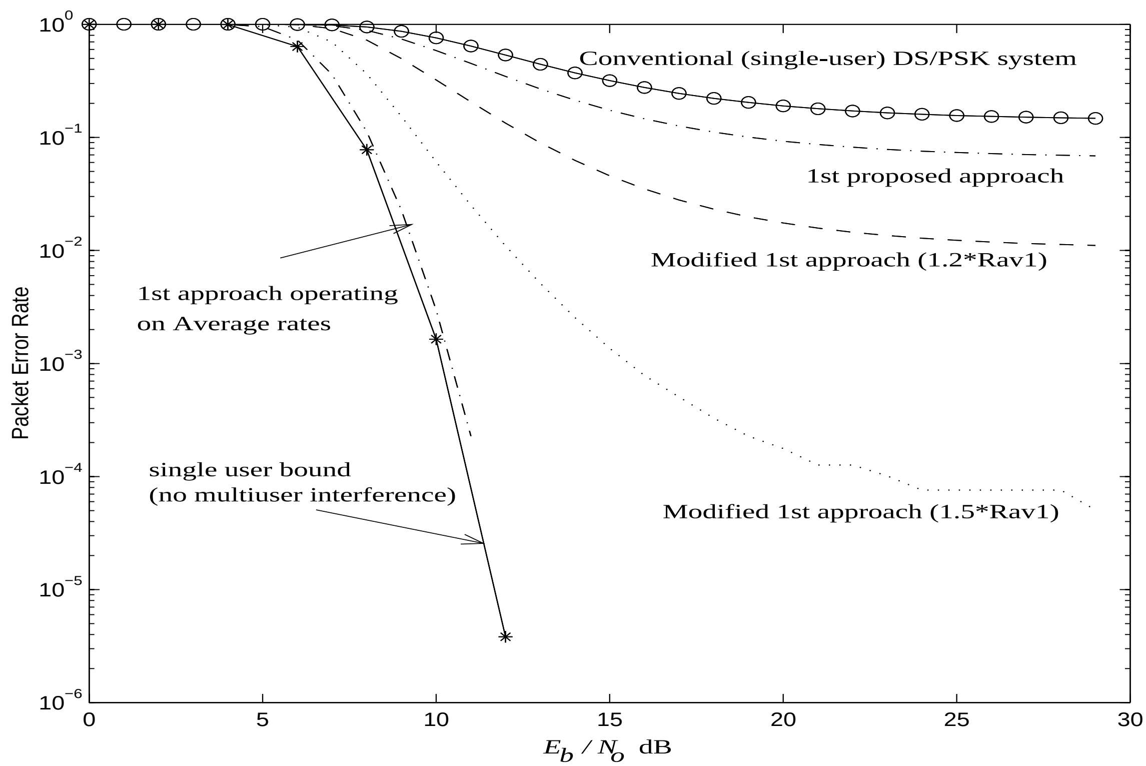


Figure 5.6: The packet error performance for 1st approach operating on high rates combinations with 60 active stream users and 13 active interactive users;$\theta^{Int}_{1,on} = 0.2, \theta^{Int}_{2,on} = 0.5, BW = 10MHz, PG = 512.$

Table 5.1 summarizes the traffic characteristics and system parameters used in evaluating the proposed traffic control approaches. The basic rate is assumed the same for all classes of both traffic categories that is $R_b = 10$ Kbps. These parameters (i.e. transmission rates) are similar to the parameters expected for the future wireless networks [25] with some little variations such that they can be used within our current system limitations. Moreover, power control and spreading code acquisition are assumed to be perfect.

Considering the stream traffic classes, we let each class has a range of transmission rates such that the proposed system will be tested under more reliable parameters. Therefore, each stream traffic class has a minimum bit rate, maximum bit rate and an average bit rate. For the purpose of investigating the capacity of

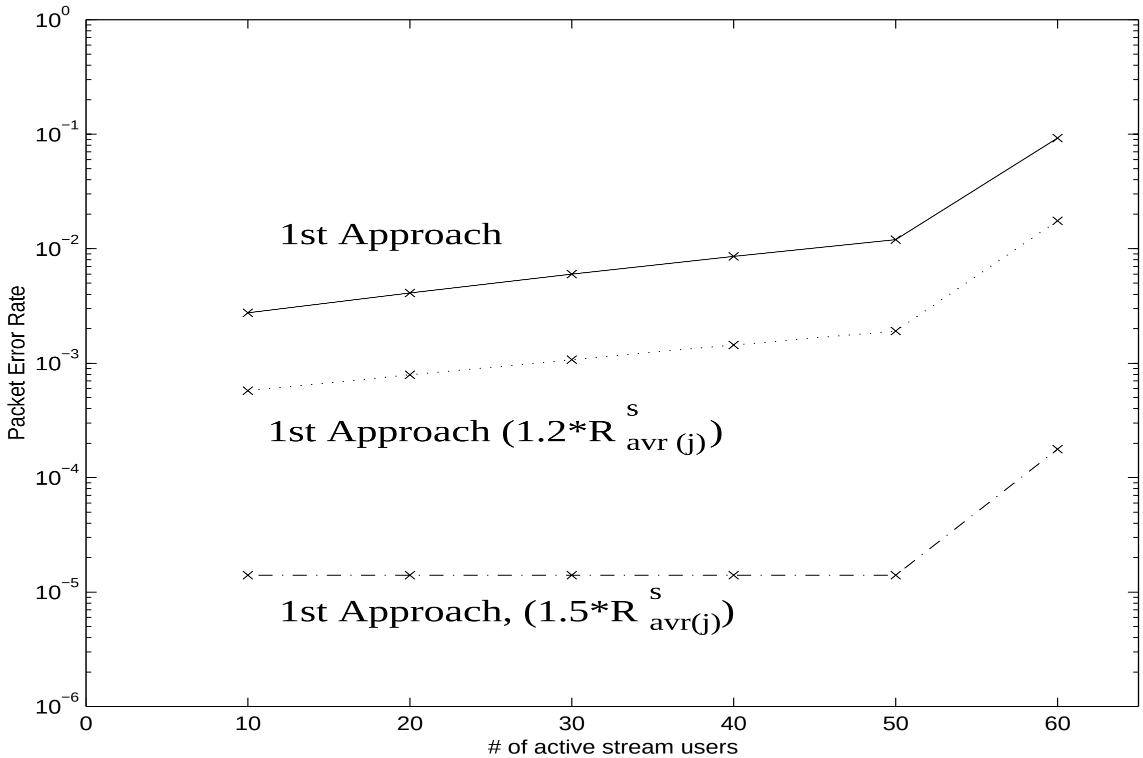


Figure 5.7: The comparison of packet error performance for 1st approach operating on high rates combinations; $BW = 10MHz, PG = 512, E_b/N_o = 20\,dB$.

the proposed protocol, we study the new system subjected to what we call combination of rates. In other words, the minimum-bit rate combination is where the system is tested assuming that all classes are operating on their lowest bit rates. The maximum-bit rate combination is where the system is tested assuming that all classes are operating on their maximum bit rates and the same thing applies for the average rates. Of course, in practice, the traffic flow is composed of different combination of transmission rates and hence, the system performance should fall within these bounds. The system studied, here, has a maximum of 60 stream users and 13 interactive users. Both proposed approaches have been studied under the above traffic characteristics and assumptions.

Figure 5.6 shows the packet error rate for the *1st* proposed approach as well as the conventional and the single user systems. It is clear that the packet error

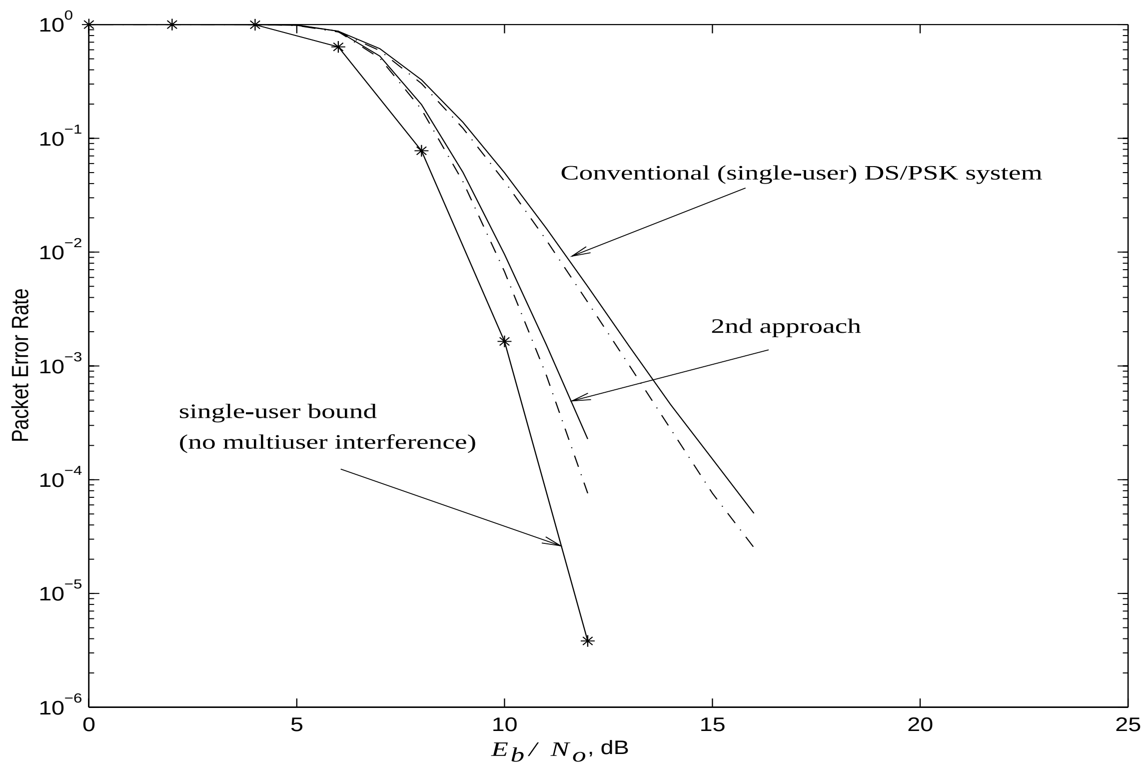


Figure 5.8: The comparison of packet error performance for 2st approach operating on minimum rates combinations with 60 active stream users and 13 active interactive users; $BW = 10MHz, PG = 512$.solid line: $\theta_{1,on}^{Int} = 0.2, \theta_{2,on}^{Int} = 0.5$, dashed line: $\theta_{1,on}^{Int} = 0.035, \theta_{2,on}^{Int} = 0.35$

rate performance is not much improved comparing to the conventional one which is irreducible. This is due to the fact that in the original proposed approach, the user should stick to a certain average bit rate and as such at high rate transmission, the packets received by the multiuser detector is just a small portion of the stream traffic. Hence, when we average the performance of both receivers over the whole traffic population, the packets received by the conventional receiver will have more statistical weight (binomial distribution). To overcome this problem, we modify the approach such that more packets shall be received by the decorrelator multiuser receiver, and this can be accomplished as follows. If the system performance is worsened by employing the original version of the proposed approach, then both the user and the base station negotiate again a new average bit rate which should

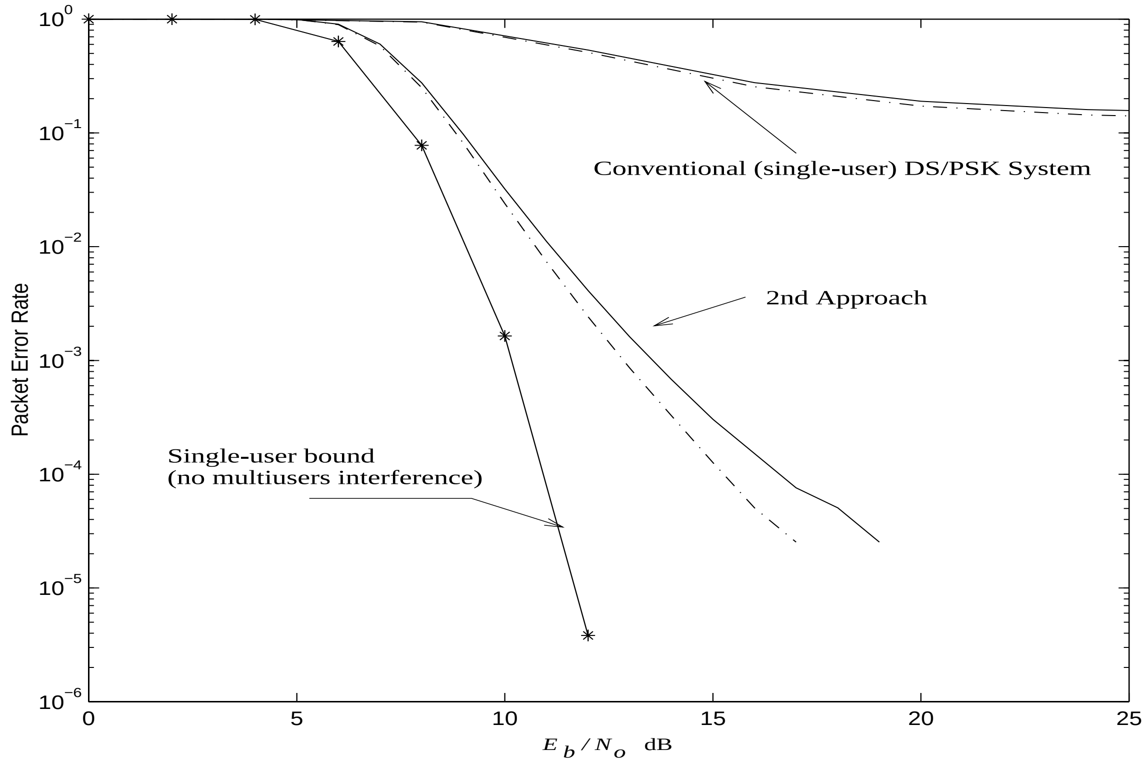


Figure 5.9: The comparison of packet error performance for 2st approach operating on maximum rates combinations with 60 active stream users and 13 active interactive users; $BW = 10MHz, PG = 512$.
solid line: $\theta^{Int}_{1,on} = 0.2, \theta^{Int}_{2,on} = 0.5$, dashed line: $\theta^{Int}_{1,on} = 0.035, \theta^{Int}_{2,on} = 0.35$

be higher than previous average bit rate (for example, $1.2R^{s}_{avr(j)}$) and so on. This modification is ,of course, limited by the practical capability of the BS (i.e. # of available receivers, etc.). Figure **??**, shows also the improvement in the system performance when this modification is applied. Further, the system capacity (i.e. the number of users the system can handle in an acceptable packet error rate) also has been improved as shown in Fig. 5.7. In this approach, the results for low-rate combinations are not shown because they are the same as for the system operating on average-rate combinations at which the *1st* approach shows very good results. On the other hand, this approach does not efficiently utilize the system capacity when the flow of traffic is lower than the average rates agreed upon at the beginning of the call.

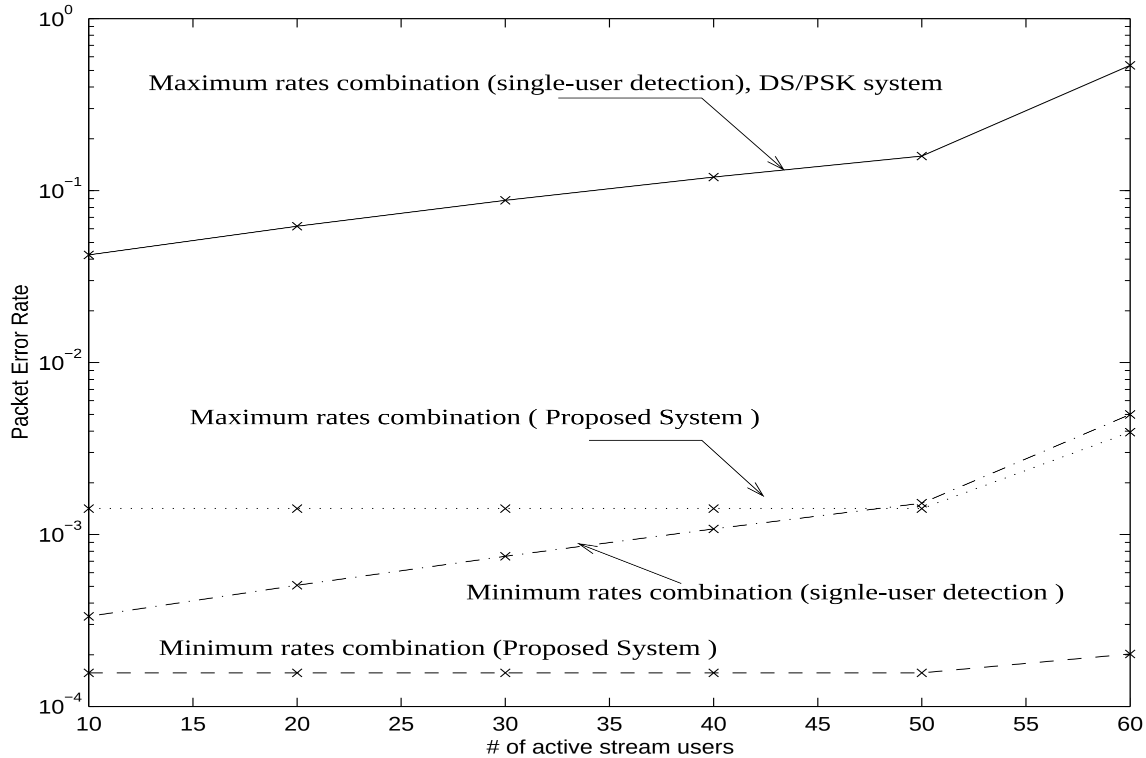


Figure 5.10: The comparison of packet error performance vs. # of active stream users for 2st approach; 13 active interactive users; $BW = 10MHz, PG = 512, E_b/N_o = 12\,dB$, $\theta^{Int}_{1,on} = 0.2, \theta^{Int}_{2,on} = 0.5$.

Now, we consider the *2nd* approach where N^s_{avr} is the statistical average estimated by the base station. Figures 5.8 & 5.9 show the comparison of the packet error rate performance of both the conventional and proposed system for the extreme cases under different burstiness activities of interactive users. It is easy to figure out the superiority of the 2nd proposed approach over all the ranges of traffic. The burstiness in the interactive traffic has an influence on the performance of the system, but it is minor which can be explained as most of the transmitted stream packets were processed by the decorrelator Multiuser receiver. Further, the 2nd proposed approach is less sensitive to the changes in the traffic parameters of stream users compared to the conventional system and this due to two factors: the robustness of the multiuser detector and the proposed traffic control that makes

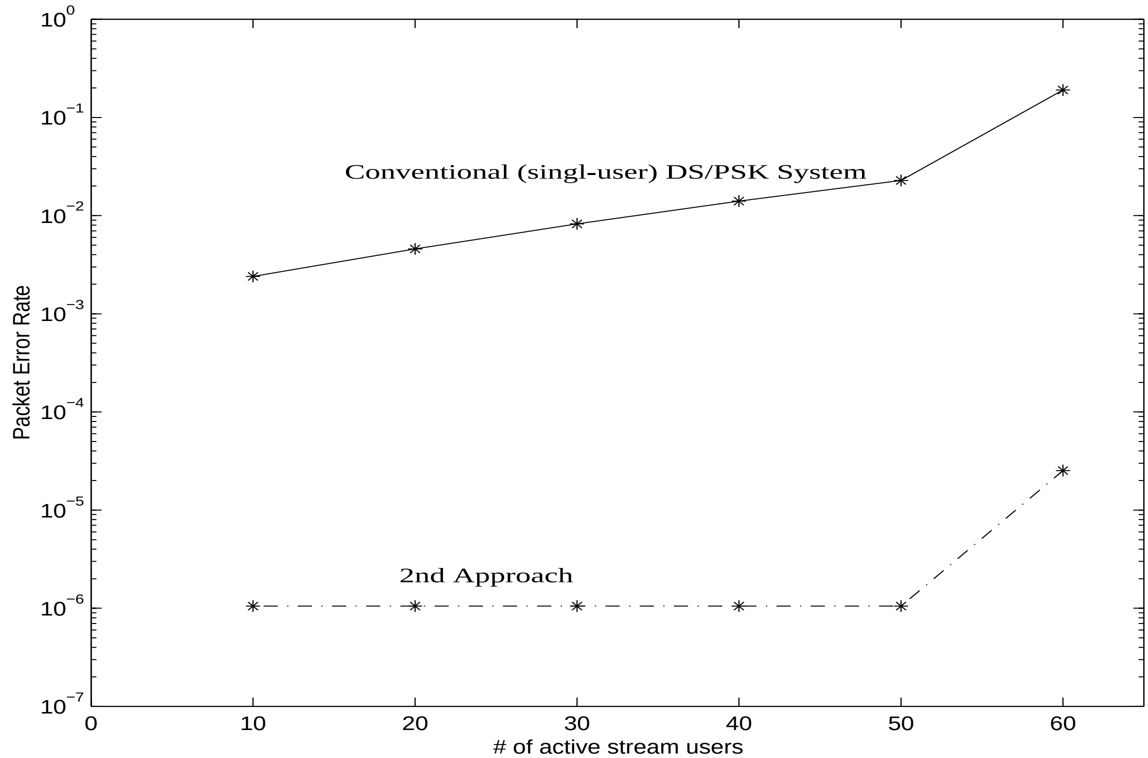


Figure 5.11: The comparison of packet error performance vs. # of active stream users for 2st approach operating on maximum rates combinations; 13 active interactive users; $BW = 10MHz, PG = 512, E_b/N_o = 20\,dB$, $\theta_{1,on}^{Int} = 0.2, \theta_{2,on}^{Int} = 0.5$.

the implementation of multiuser receiver more practical and be able to balance the flow of traffic into two different detection strategies and achieve a good compromise performance. In fact, this insensitivity of the proposed system to traffic changes is directly translated into a larger system capacity compared with the conventional system. This last notion is very obvious in Figs. 5.10 & 5.11. Nevertheless, as contrary to the first approach, the second approach highly depends in its operation on the idealism of the BS estimation of the average number of transmitted packets from all active stream users on the first slot (i.e. T_{s1}). Figure 5.12 compares the achieved system capacity by both proposed approaches.

It is intuitive that the more packets received by the multiuser receiver, the

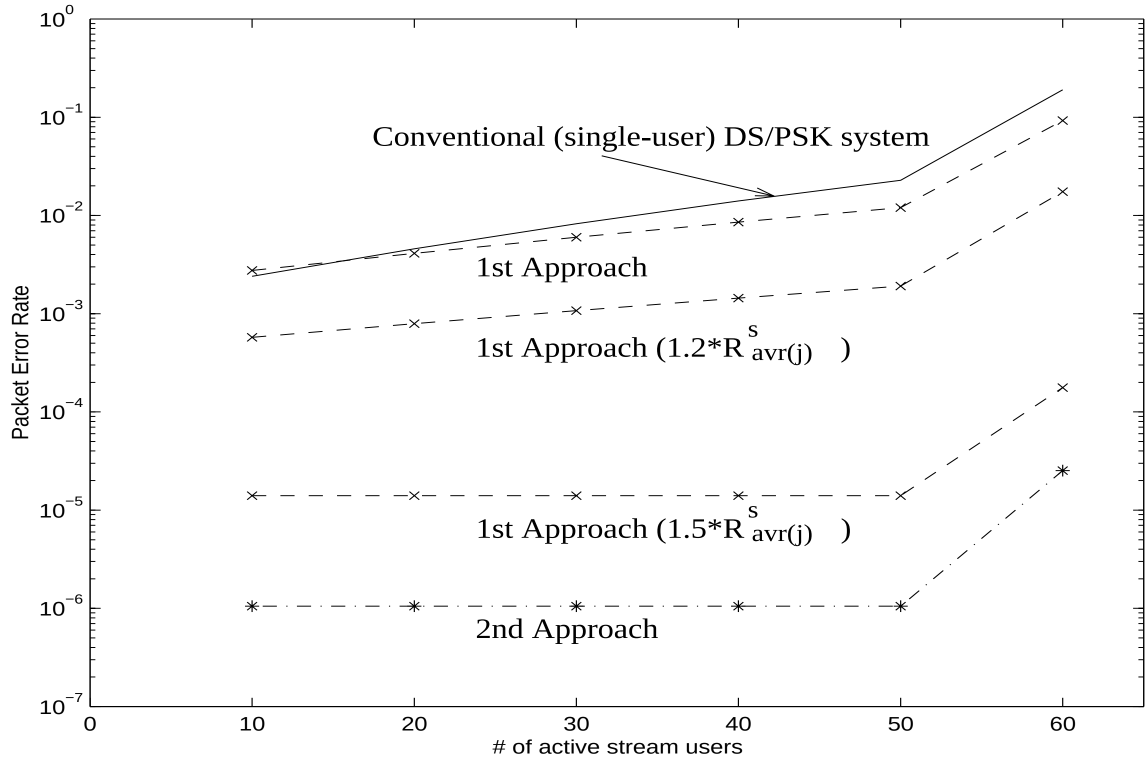


Figure 5.12: The comparison of packet error performance for both approaches operating on maximum rates combinations; 60 active stream users, 13 active interactive users; $BW = 10MHz, PG = 512$, $\theta^{Int}_{1,on} = 0.2, \theta^{Int}_{2,on} = 0.5$.

better the receiver performance is achieved. However, the price paid is the complexity of the receiver is enhanced as well as the expected delay the transmitted packets might suffer which would become a real problem for delay-sensitive applications such as voice, video users. Hence, our proposals try to balance the traffic between the two time slots such that each time slot can achieve an acceptable performance. To investigate more on this point, we change the ratio of $\frac{N^s_{avr}}{N^s_{inst}}$ in the stream traffic population and by such we are varying the number of packets emitted from stream traffic users to each time slot. Figure 5.13 shows the effect of changing the ratio of the packets processed by the multiuser receiver to the total number of packets emitted from the stream traffic on the packet error rate performance. Finally, Fig. 5.14 compares the packets error rate performance of both categories of traffic (i.e.

stream and interactive). We notice that even the proposed system may enhance the burden on T_{s2}, the performance of the interactive traffic in T_{s2} is almost the same as the stream traffic for the same E_b/N_o.

5.6 Conclusions

A new hybrid TDMA/MC-CDMA system accompanied with two traffic flow control approaches has been presented. The practicality and bit error performance of the new system have been examined under a wide range of expected traffic characteristics (bit rate, transmission activities, etc.) for the future wireless networks. The results show the superiority of the proposed system compared with the conventional one. Further, this improvement in the performance is attributed to the novel traffic control approaches that have introduced an interaction between the physical layer and higher network layers and consequently balance the traffic load on each detection algorithm and makes the implementation of the decorrelator receiver more practical.

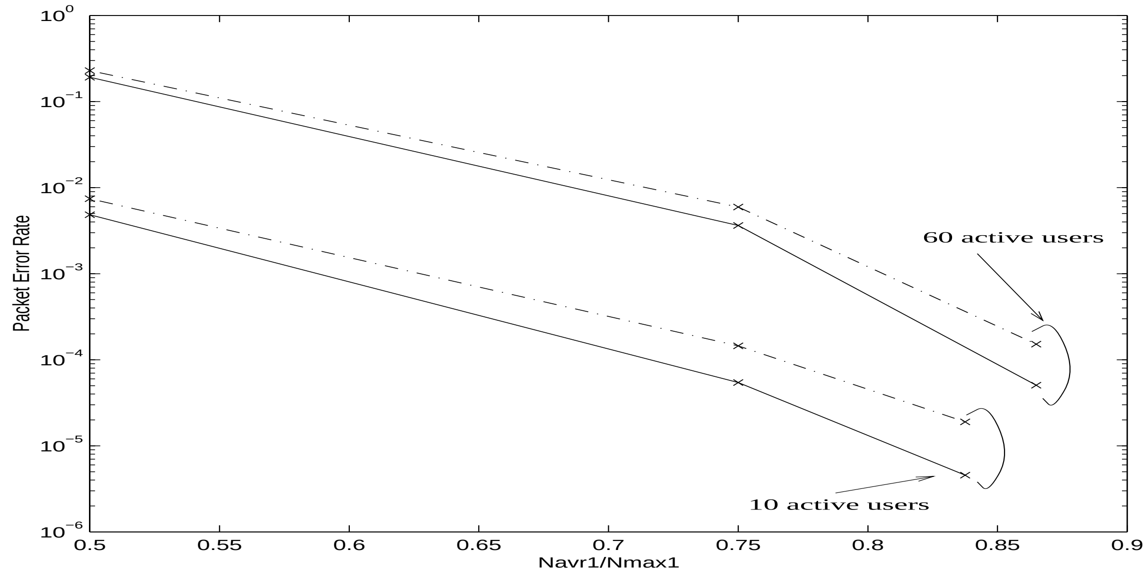


Figure 5.13: The effect of the changes in $\frac{N^s_{avr}}{N^s_{inst}}$ ration on the packet error performance for the proposed systems operating on maximum rates combinations; 13 active interactive users; $BW = 10MHz, PG = 512, E_b/N_o = 16\,dB$, solid line: $\theta^{Int}_{1,on} = 0.2, \theta^{Int}_{2,on} = 0.5$, dashed line: $\theta^{Int}_{1,on} = 0.035, \theta^{Int}_{2,on} = 0.35$.

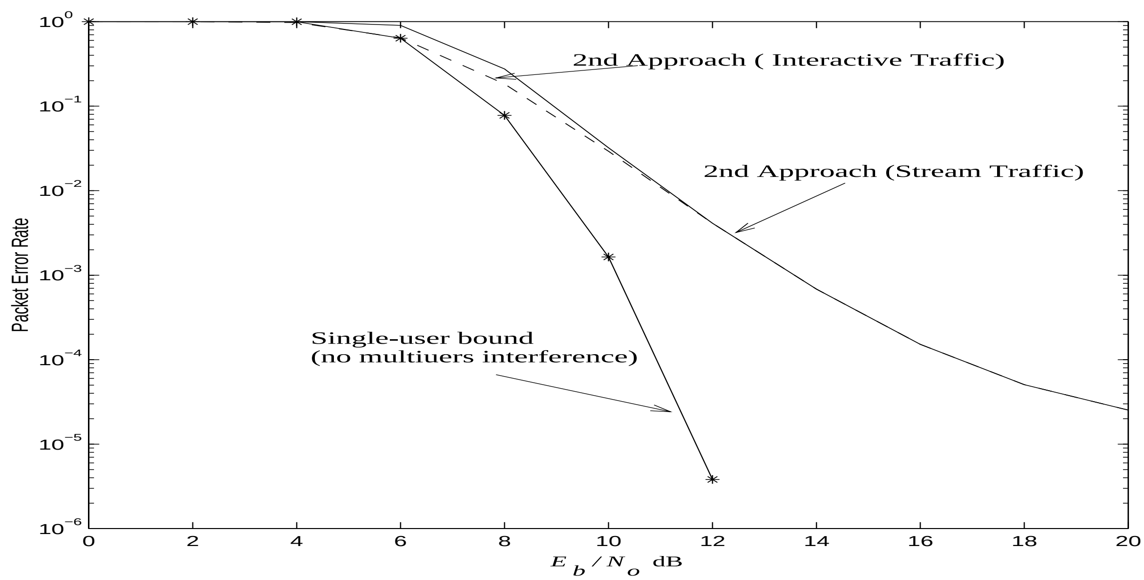


Figure 5.14: The comparison of packet error performance for 2st approach operating on maximum rates combinations; 60 active stream users, 13 active interactive users; $BW = 10MHz, PG = 512$, $\theta^{Int}_{1,on} = 0.2, \theta^{Int}_{2,on} = 0.5$

Chapter 6

Hybrid MC-CDMA/TDMA: Queueing Analysis

6.1 Introduction

In the preceding chapter, we discussed two flow control approaches applied to a hybrid CDMA/TDMA platform. The results shows that the proposed approaches have outperformed the conventional CDMA system. These results, however, are obtained under ideal conditions in terms of resources (e.g. receivers, transmitters, buffer size, etc.) availability at the mobile stations as well as the base station. In practice, there are limitations for these resources and we should expect "network problem issues" such as call blocking, cell blocking, delay, delay jitter, etc. Therefore, these proposed approaches should be accompanied with admission policy which reflects on these limitations at any admission decisions.

In this chapter, we develop a general queueing model for the end-to-end performance analysis of CDMA networks. In this model, we interrelate the physical limitations of the base stations (i.e., the number of transmission and reception modems), call and burst level traffic, instantaneous ATM buffer conditions, and End-to-End bit error performance in one queueing problem. Previous studies do not include the error performance in the queueing modeling of the CDMA networks nor both uplink and downlink performance in one problem. Nevertheless, the inclusion of the channel performance in the queueing modeling would give more realistic figures of network performance as well as guarantees of the required QoS of the networks. First, an algorithm is developed for a general queue and then another algorithm is presented for general priority scheme that easily can be applied along with the general algorithm.

Moreover, the state diagram representing such kind of queueing problem is so general that it does not fit the description of typical queues. Therefore, analytic solution by means of moment generating function is intractable and only solution of the $(B + 1)$ simultaneous equations (where B is the base station total ATM buffer capacity in cells) describing the probability balance equations of each state will be pursued.

6.2 Queueing Analysis

In all analysis presented in this chapter, the considered system has a maximum of L servers on the downlink channel and a finite buffer capacity of B cells (packets) (usually, $L \leq B$). Further, there is no restriction on the arrival process nor on the service process distributions. These distributions can be implicitly included in the transition probabilities among the buffer states. In the following, we will develop algorithms for calculating the transition probabilities. Also, in this section, we do not restrict the analysis to our MC-CDMA/TDMA proposal and it is presented in a general way which can be applied to any multiple access protocol. Later, however, we shall restrict the analysis to study our MC-CDMA/TDMA proposal.

6.2.1 Priority Queues

The diversity of QoS parameters of future networks users requires that the base station be capable of granting users different levels of priority such that these QoS requirements are guaranteed. We assume that at the beginning of the call, the user is assigned certain priority level according to which class of traffic he belongs to. For simplicity of presentation, users are assigned priority in ascending order (i.e. the highest priority user is assigned 1). Further, each class of users has its own queue at the base station. The following algorithm is general for calculating the transition probabilities of any group of priority queues.

Considering the kth class of traffic, we define ${i^k}_p$ as the instantaneous number of cells belonging to all traffic classes that have been assigned higher priority, i.e.

$$ {i^k}_p = \sum_{n=1}^{k-1} j_n \tag{6.1} $$

where j_n is the instantaneous generated cells from an nth priority class of traffic. Also, define i_k as the total number of instantaneous cells belonging to the kth class of traffic as well as all classes of traffic that are assigned higher priorities, i.e. $i_k =$

$j_k + {i^k}_p$. Define $\gamma_j = \lambda_j u_j$ as the probability that j cells have been generated with probability λ_j which is dependent on the probability distribution of the nth class of traffic and received successfully with probability u_j at the base station. similarly, we define d_j as the probability of successfully transmitted cells via the downlink channel. Clearly, u_j and d_j are functions of modulation/demodulation, coding/encoding, uplink/downlink environments, etc. Since there is a maximum of L servers on the downlink, the maximum allowed cells on the downlink will be L. Therefore, we define $inst$ as the instantaneous number of cells generated from all active uses that are allowed to share the downlink channel simultaneously, i.e. $inst = min(L, \sum_{n=1}^{K} j_n)$, where K is the total number of classes of traffic population.

- *for* $i < j \leq B$ (birth-process)

  for $i_k > L$

$$p_{ij} = \begin{cases} \gamma_{j-i_k+L} d_L + \gamma_{j-i}(1 - d_L) & ; {i^k}_p \leq L \\ \gamma_{j-i} & ; {i^k}_p > L \end{cases} \tag{6.2}$$

  for $i_k \leq L$

$$p_{ij} = \gamma_j d_{inst} + \gamma_{j-i}(1 - d_{inst}) \tag{6.3}$$

- *for* $B \geq i > j$ (departure-process)

  for $(i - j) \leq L - {i^k}_p$

$$p_{ij} = \begin{cases} \gamma_j d_{inst} & i_k \leq L \\ \gamma_{j-i_k+L} d_L & i_k > L \end{cases} \tag{6.4}$$

6.2.2 General Queues

Now, we consider a general structure of a buffer that its transition probabilities interrelate the uplink and down link error performance, the uplink receivers, the down link servers as well as the buffer capacity. The arrival process is random while the service process is deterministic. Here, we present the algorithm for calculating the transition probabilities for such queue. Figure 6.1 illustrates the state diagram for a simple example of such queue. It is easy to see that one may impose any

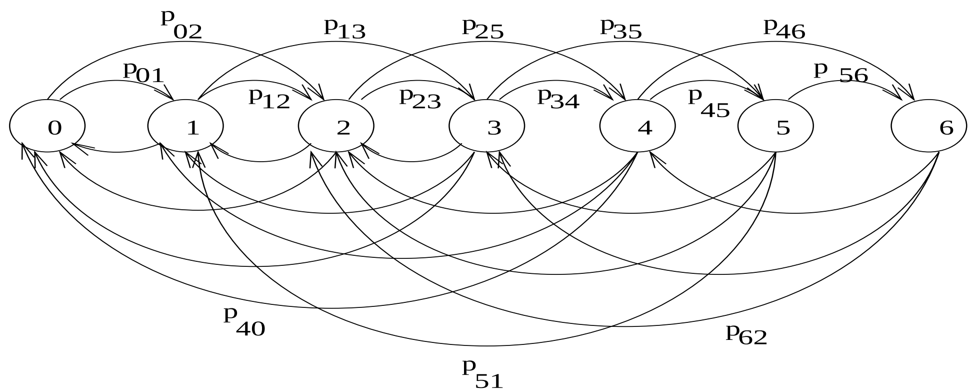


Figure 6.1: Illustration of the transition probabilities of the traffic; L=4, B=6, Ex=2 cells.

priority scheme on these traffic population by just running this algorithm under the priority queue algorithm mentioned in section 6.2.1.

Defining Ex as the maximum possible received cells during one cell-time unit, the transition probabilities of the state diagram are given by:

- *for* $i < j \leq B$ (birth-process)

  for $j - i < Ex$ or $j = Ex$

$$p_{ij} = \begin{cases} \gamma_{j-i+L} d_L + \gamma_{j-i}(1 - d_L) & ; i > L \\ \gamma_j d_i + \gamma_{j-i}(1 - d_i) & ; i \leq L \end{cases} \tag{6.5}$$

  for $j - i = Ex$

$$p_{ij} = \begin{cases} \gamma_{j-i}(1 - d_L) & ; i > L \\ \gamma_{j-i}(1 - d_i) & ; i \leq L \end{cases} \tag{6.6}$$

  $p_{ij} = 0$; Otherwise

- *for* $B \geq i > j$ (departure-process)

  for $L \geq i$

$$p_{ij} = \begin{cases} \gamma_j d_i & ; j \leq Ex \leq L \text{ or } Ex > L \\ 0 & ; \text{Otherwise} \end{cases} \tag{6.7}$$

for $L < i \leq L + j$

$$p_{ij} = \begin{cases} \gamma_{j-i+L} d_L & ; j - i + L \leq Ex \leq L \text{ or } Ex > L \\ 0 & ; \text{Otherwise} \end{cases} \tag{6.8}$$

- *for* $i > j$ and $i - j > L$

  $P_{ij} = 0$

It is interesting to note that the state diagram developed by the above algorithm consider the content of cells of the buffer regardless of the sources of these cells. In other words, if a buffer has j cells, then, these cells are not necessarily belonging to one traffic source. They might arrive from several traffic sources. This means that these state diagrams are multidimensional. This multidimensionality is implicitly imposed by the probability transitions. Further, this generality in representing the state diagrams manifests its self by the presence of transition to distant states by the limiting of service to a maximum of only L simultaneous cells per cell time on the down link channel and by the dependence on successful transmission probability via the downlink channel d_i. It does not fit the description of typical queues. Analytic solution by means of moment generating function is not convenient for this general queue and only solution of the $(B+1)$ simultaneous equations (where B is the base station total ATM buffer capacity in cells) describing the probability balance equations of each state will be pursued.

Writing the balance equations of the steady-state transition probabilities of the state diagram, we get a set of linear equations that can be formulated in a matrix form as

$$A\underline{X} = \underline{Y} \tag{6.9}$$

where $\underline{X} = [X_0 X_1 \cdots X_B]^T$, $\underline{Y} = [Y_0 Y_1 \cdots Y_B]^T$ and the elements of$A$ can be obtained as in (6.10)

$$a_{ij} = \begin{cases} -\sum_{\substack{k=i_{min} \\ i \neq k}}^{B} p_{ik} & ; i = j \\ p_{ji} & ; i \neq j \end{cases} \tag{6.10}$$

where i_{min} is the least state in the given state diagram. For example, i_{min} is the average number of cells if the average stream queue is considered, while i_{min} is zero if the collective traffic queue is considered. This set of equations is typically complemented by the condition that

$$X_0 + X_1 + X_2 + \cdots + X_B = 1 \tag{6.11}$$

Now, by replacing the last row in A with the all ones row (corresponding to equation (6.11)), the steady state probabilities X_l's will be found by solving the system of $(B+1)$ simultaneous equations in (6.12).

$$\begin{bmatrix} a_{00} & a_{01} & \cdots & a_{0B} \\ a_{10} & a_{11} & \cdots & a_{1B} \\ a_{20} & a_{21} & \cdots & a_{2B} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} X_0 \\ X_1 \\ \vdots \\ \vdots \\ X_B \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 1 \end{bmatrix} \tag{6.12}$$

Two-Queue Scenario:

Now, we apply the above algorithms on the proposed MC-CDMA/TDMA protocol. We assume that the population of each of the 2 TDMA slots has its own buffer at the base station. Thus, the average traffic of stream users are queued in one buffer while the excess traffic as well as any cells from interactive traffic (if it is allowed) are queued in the second buffer. Therefore, each queue should be analyzed differently. Figures 6.2 illustrates the state diagrams for simple example of the stream traffic transmitted via **Slot-I**.

Considering the average stream traffic queue, we are certain that during **Slot-I**, there are Av cells transmitted and detected by the multiuser receiver, where Av is the pre-negotiated average cells. Hence, this queue has deterministic arrival as well as deterministic service processes and we shall denote it as $D/D/L/B$ queue. The probability transitions of the underlining state diagram are as follows.

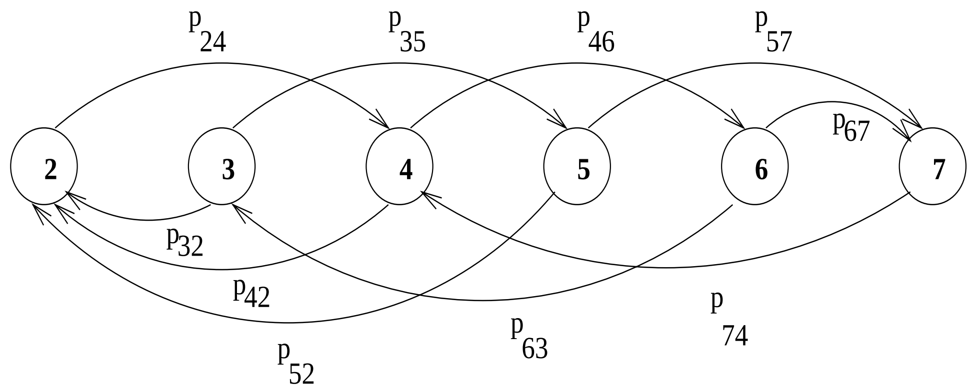


Figure 6.2: Illustration of the average stream traffic queue; $L = 5, B = 7, Av = 2$ cells.

- *for* $L > Av$

  for $i < j$, where $i \geq Av, j = min(i + Av, B)$ (birth-process)

$$p_{ij} = \begin{cases} u_{Av}(1 - d_i) & i \leq L \\ u_{Av}(1 - d_L) & i > L \end{cases} \tag{6.13}$$

  for $B \geq i > j$, where $j = i - L + Av$ (departure-process)

$$p_{ij} = \begin{cases} u_{Av}d_i & i \leq L \\ u_{Av}d_L & i > L \end{cases} \tag{6.14}$$

- *for* $L \leq Av$, the queue is unstable

On the other hand, the queue that receives the transmitted cells via **Slot-II** have completely different characteristics. The arrival process is random while the service process is deterministic. Here, the transition probabilities can be calculated directly applying the above algorithms. More, Fig. 6.1 can also be used as an illustration for the packets transmitted using **Slot-II** assuming that Ex is the maximum possible excess packets received during one slot.

In the succeeding chapter, we will use these queueing models to determine the steady state probabilities of buffer occupancy at the base station. Then, these probabilities will play a key role in our admission/congestion policy that shall applied to the hybrid CDMA/TDMA network under the first proposed flow control (see section 5.2.1).

Chapter 7

Window Measurement-Based Admission/Congestion Control

7.1 Introduction

Having developed algorithms for calculation of transition probabilities for general priority queues in chapter 6, we are ready to apply these algorithms to evaluate CDMA networks for several performance measures in different scenarios. We shall start with introducing a new adaptive call admission policy based on a Window-Measurement estimation of the status of the Queue at the base station. Then, we first analyze and evaluate the performance of the so called 'pure' CDMA network under the new admission policy. Next, we look at the performance of the proposed MC-CDMA/TDMA protocol under the new admission policy.

7.2 Admission/Congestion Control Policies

The proposed admission/congestion policy has two phases. Phase-I (admission) manages the admission of new calls such that there are enough resources (servers) on the uplink to serve these new users. The proposed traffic flow control mentioned above imposes the following restriction. Stream users should be admitted into the system, if the available uplink servers (maximum of L^u) can accommodate the variable rate components (excess traffic) of active stream users concurrently with the active interactive users, i.e.,

$$N^s_{excess} + N^I \leq L^u \tag{7.1}$$

where N^I is the instantaneous number of cells generated from active interactive users. This means that the admission policy grants higher priority to interactive users calls which is very practical for delay-sensitive application such as voice, video calls. When Eq. (**??**) is not satisfied, *contention* is identified in this thesis. On the the hand, we can assume the same level of accessibility for both traffic categories providing that the stream users should restrict their transmission rates to the average rates if all servers in slot-II are busy serving the interactive users. This modification would lead to a larger delay especially for high variable rate users.

Secondly, phase-II (congestion) of the proposed traffic control policy manages the cell flow into the base station such that it prevents any possible congestion in the buffer. This objective is achieved by interrelating the End-to-End network parameters such as up/down link channels, number of receivers at the base station, buffer limitation at base station, traffic burstiness, etc. in one *Queueing Problem*. Subsequently, the instantaneous buffer overflow at the base station is used to predict the likelihood of buffer congestion over a period of a window of transmitted packets. This reactive congestion policy is proposed and examined in the following sections.

7.2.1 Window Measurement-Based Admission Policy

Figure 7.1 shows an ensample of a network activity at the base station. We denote the case when the number of cells in the buffer exceeds certain threshold TH by an upward arrow. Otherwise, it is denoted by a downward arrow and this what we call the buffer is in a 'good' condition. This randomness is due to many factors, such as the burstiness during the cell level, the erroneous channels, etc. These activities also vary from window to window due to the above reasons as well as the call activity of each user.

Using the occupancy of the buffer, we define the cell overflow (or loss) probability $O_b(.)$ such that it gives an early warning about the build up process of queued cells in the system.

$$O_b(.) = \sum_{l=TH}^{B} X_l \tag{7.2}$$

where TH is the buffer content that defines congestion. Equation (7.2) defines buffer overflow over one cell time. The buffer overflow probability is an important indicator of how the overall system is working. As was previously stated, maintaining the QoS requirements is a crucial issue for any integrated services wireless network. In particular, blocking probability and cell delay are interrelated entities. So developing an estimator that can estimate and predict the likelihood of buffer congestion, helps

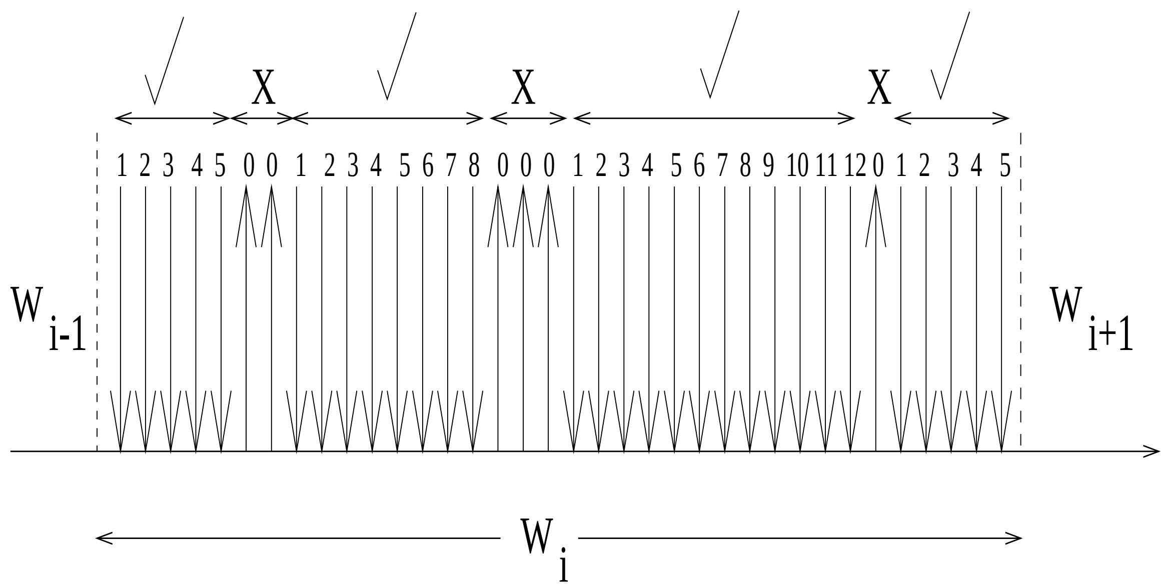


Figure 7.1: Illustration of buffer activities at the BS; window size= 36 cells.

to put these related and crucial performance issues (i.e. delay, blocking probability) under control.

If the dialup or call signaling period equals W cells, the probability that an incoming call finds the buffer in good condition for at least q consecutive cells within these W cells (observation or dialup period), which we call probability of buffer not congested $S(.)$ can be shown (see Appendix B) to be:

$$\begin{aligned} S(.) \quad \leq \quad & \sum_{i=0}^{W-q-2} \begin{pmatrix} W \\ i \end{pmatrix} (O_b)^i (1-O_b)^{W-i} + (W-q)^2 (O_b)^{W-q-1} (1-O_b)^{q+1} \\ & + (W-q+1)(O_b)^{W-q}(1-O_b)^q \end{aligned} \tag{7.3}$$

$S(.)$ will be averaged over all possible combinations of admitted calls, i.e.

$$\overline{S} = \sum \cdots \sum S(.)\, C(.) \tag{7.4}$$

where $C(.)$ is the joint probability distribution of a certain combination of active users initiated from all traffic categories. Later, we will be more specific about $C(.)$ which depends on the traffic probability distribution.

It is obvious that the probability of non-congested buffer reflects the dynamics of the building up process of the queued cells during the call period. Hence, considering this measure as an acceptance or rejection criterion is warranted. Consequently, the traffic load must be adapted according to $\overline{S}$ as shown in (7.5).

$$\rho_{new} = \rho \cdot \overline{S} \tag{7.5}$$

Where ρ_{new} is the average accepted traffic, and ρ is the total stream offered traffic load (Erlangs). According to (7.5), $\overline{S}$ is actually the average throughput of the system. For example, if $\overline{S} = 0.9$, then only 90% of the traffic should be admitted and the rest must be rejected. By controlling the traffic load as in Eq. (7.5), the following quality of service parameters, namely, cell blocking probability call establishment delay and the cell error will be under control. Therefore, the second phase of the proposed policy (congestion) has two steps. First the likelihood of buffer congestion ($\overline{S}$) is estimated. Second, the offered traffic load is adapted accordingly. This adaptation of the offered traffic can be implemented in different ways. For example, the transmission rates of all currently admitted users can be varied such that the accepted traffic load matches with $\overline{S}$ and the number of admitted calls is kept without change. On the other hand, we can choose to take an opposite strategy that is to keep the transmission rates as they are and instead block some calls following certain policy.

Now we are in a position to state our new call admission policy; i.e., the new admission states "The new call is admitted once the total number of low-bit streams corresponding to the active calls is less than L^u as in equation (7.1) and the buffer is not congested as in equation (B.7)." The overall new call blocking probability in this case is given by

$$Bl_2 = P_{block} = 1 - \overline{S} \tag{7.6}$$

We note that the call admission policy (i.e. phase I) leading to Bl_1 is easier to implement since it involves no direct ATM cell buffer measurement (it only monitor the number of calls), while the new policy leading to Bl_2 necessitates monitoring the number of cells in the total buffer of the base station,on top of monitoring cells. This number varies from slot to slot (one cell time on the downlink channel), and if actual buffer measurements are based on one cell time, $Bl2$ will be excessively high since a spontaneous congestion (one cell time) will lead to new calls blocking (i.e., an overreaction). On the other hand, if the measurement time underlining Bl_2 is excessively long (e.g., few cell bursts) this smoothes down the congestion, more new calls will be admitted but the buffer may experience high level of congestion ($X_i \geq TH$) for short periods of time (call bursts of part thereof) thus increasing the user delay jitter, In this work the ATM measurement window, W, is assumed to be of the order of the smallest average call active burst length. Now, it is easy to find the mean call establishment time, i.e.

$$\begin{aligned} \overline{CE^s} &= \textstyle\sum_{l=1}^{\infty} l \cdot (P^s_{block})^{l-1}(1 - P^s_{block}) \cdot (W + \tau) \\ &= \tfrac{(W+\tau)}{(1-P^s_{block})} = \tfrac{(W+\tau)}{\overline{S^s}} \qquad \text{cell times} \end{aligned} \tag{7.7}$$

where τ is the round trip propagation delay between user and BS (or satellite), involved in each user signaling trial of duration W cells.

7.3 Statistical Traffic Models

Considering that I different classes of users are multiplexed, the aggregated cell arrivals are governed by the number of calls in the ON state from all types of active users. Having assumed that all classes of traffic users are independent identically distributed random variables (iid), the compound traffic distribution of i cells generated by all active calls (i.e. λ_i, $i = \sum_{k=1}^{I} n_k$) of all I classes in a certain category is given by

$$\lambda_i = \sum_{(n_1,\cdots,n_I)\in\mathcal{R}} \prod_{k=1}^{I} P_k(n_k) \tag{7.8}$$

where

$$\mathcal{R} = \left\{ (n_1, \cdots, n_I) \quad : \sum_{k=1}^{I} n_k = i, \quad n_k \leq N_k \forall k \right\}$$

where N_k is the maximum possible number of cells generated by class k and $P_k(n_k)$ is the probability distribution of generating n_k cells from class j which is given by

$$P_k(n_k) = \left(\begin{array}{c} N_k \\ n_k \end{array} \right) (\theta_{k,on})^{n_k} (1 - \theta_{k,on})^{N_k - n_k} \tag{7.9}$$

Similarly, at the call level, we assume that all traffic sources are modeled by the ON/OFF model with different activity factor for each class of traffic. Therefore, the joint probability distribution of j admitted calls (i.e. $\Gamma_j, \ j = \sum_{k=1}^{J} j_k$) from J different classes of users is given by

$$\Gamma_j = \sum_{(n_1, \cdots, n_I) \in \mathcal{R}} \prod_{k=1}^{J} Q_k(j_k) \tag{7.10}$$

where

$$\mathcal{R} = \left\{ (n_1, \cdots, n_I) \quad : \sum_{j=1}^{I} n_j = i, \quad n_j \leq Su_j \forall j \right\}$$

where Su_j is the number subscribers to class k and $Q_k(j_k)$ is the probability distribution of initiating j_k calls from class k which is given by

$$Q_k(j_k) = \left(\begin{array}{c} Su_k \\ j_k \end{array} \right) (\delta_{k,on})^{j_k} (1 - \delta_{k,on})^{Su_k - j_k} \tag{7.11}$$

where $\delta_{k,on}$ is the activity factor of class k users (i.e. call/sec). It is necessary to note that the probability distributions of call and cell levels are presented in a general way without distinguishing between stream users and interactive users. Later in section 7.5.1, these distributions shall be identified according to their traffic categories (i.e. the superscript 's' for stream users and 'Int' for interactive users).

7.4 "Pure" CDMA Platform System performance [96]

Here, we are going to apply the above admission/congestion policy on pure CDMA platform where both uplink and downlink are operating on CDMA platform similar to **IS-95**. All traffic classes' users can access the network simultaneously. Further, we assume that all receivers in the network are single-user receivers.

In evaluating the performance of the aforementioned admission policy, we consider the following system specifications. We assume that the uplink traffic and the downlink traffic are using different channels. The source stream bits are modulated using BPSK-DS-CDMA with a processing gain of 255. The stream bits are encoded by convolutional codes. Further, following the **IS-95** specifications [99], the uplink bits are encoded by a convolutional code where $k = 9, R = \frac{1}{3}$, while the downlink bits are encoded by another convolutional code with $k = 9, R = \frac{1}{2}$. It is assumed that the cell on the uplink channel are received noncoherently at the base station. On the other hand, the transmitted cells via the downlink channel are assumed to be received coherently by the mobile stations. Both links are assumed to be under the influence of two paths Rayleigh multipath fading.

In this analysis, it is assumed that we have only two classes of traffic, namely, voice and data. Voice users have been granted higher priority and not buffered while data users have less priority and they are allowed to be queued. Therefore, voice cells experience no delay except the negligible propagation delay. Further, we assume that the average arrival of voice calls is double the average arrival of data calls (i.e. $\lambda_v = 2\lambda_d$ calls/s), while both traffic sources are generating cells at the same rate during the active period. In our results, we let $R_v = R_d = 10\ Kbps$, which can be translated into approximately 23 ATM cells/sec. However, taking into consideration the traffic characteristics of voice and data users (e.g. burstiness, call duration, etc.), we assume that $\theta_{v,a} = 0.5$ while $\theta_{d,a} = 0.8$. It is assumed that no new call is initiated by the busy user until the current call is completed and the user becomes idle again. The buffer capacity is limited to 70 cells (i.e. B=70). On the other hand, the maximum number of receivers at the base station is 64. The measurement window W equals to 150 cells, and q is taken to be 90 cells.

We should follow the same analysis developed in section 7.3. Nevertheless, we will reformulate the above analysis to match the new problem where we just have one slot on each link and only two traffic sources are considered.

7.4.1 CDMA Link Performance Analysis

By the independence of users transmission on the uplink (reverse link), we obtain the probability distribution of the total number of cells generated from all voice (j_v) and data (j_d) active calls; i.e.,

$$\begin{aligned} \lambda_i(jv, jd) &= \textstyle\sum_{k=0}^{j_v} \sum_{j=0}^{j_d} P_v(k) P_d(j) \\ P_v(k) &= \begin{pmatrix} j_v \\ k \end{pmatrix} (\theta_{v,a})^k (1-\theta_{v,a})^{j_v-k} \\ P_d(j) &= \begin{pmatrix} j_d \\ j \end{pmatrix} (\theta_{d,a})^j (1-\theta_{d,a})^{j_d-j} \end{aligned} \tag{7.12}$$

where $i = k + j$, for $i = 0, 1, \cdots, j_v + j_d$. We should notice that the dependence of all the parameters in the following equation (j_v, j_d, i) will be occasionally dropped for clarity purposes.

7.4.2 Window-Based Admission Policy

Let X_l^d is the probability of having l cells in the buffer, when j_v voice calls and j_d data calls are active. We define the cell overflow (or loss) probability O_b^d such that it gives an early warning about the build up process of queued calls in the system.

$$O_b^d(j_v, j_d) = \sum_{l=TH}^{B} X_l^d \tag{7.13}$$

where TH is the buffer content that defines congestion. The probability of buffer is given by:

$$S(j_v, j_d) \leq \sum_{i=0}^{W-q-2} \begin{pmatrix} W \\ i \end{pmatrix} (O_b^d)^i (1 - O_b^d)^{W-i}$$
$$+(W-q)^2 (O_b^d)^{W-q-1} (1 - O_b^d)^{q+1} \tag{7.14}$$
$$+(W-q+1)(O_b^d)^{W-q} (1 - O_b^d)^q$$

Then, $S^d(.)$ will be averaged over all possible combinations of admitted calls, i.e.

$$\overline{S^d} = \sum_{j_v} \sum_{j_d} S^d(j_v, j_d) C(j_v, j_d) \tag{7.15}$$

where $C(.)$ is the joint probability distribution of a certain combination of active users initiated from both traffic sources (i.e. voice and data). Assume for simplicity Poisson arrivals of calls for voice and data with total calls arrival rates λ_v and λ_d, and holding times $\frac{1}{\mu_v}$ and $\frac{1}{\mu_D}$, respectively. Therefore, the probability of having j_v voice calls and j_d calls on the uplink channel (assuming an $M/M/1$ model for the calls activities) is Geometrically distributed, i.e.

$$C(j_v, j_d) = (1 - \rho_v) {\rho_v}^{j_v} \cdot (1 - \rho_d) {\rho_d}^{j_d} \tag{7.16}$$

Therefore, the traffic load must be adapted according to $\overline{S^d}$ (i.e. $\rho_d^{new} = \rho_d \cdot \overline{S^d}$, where ρ_d^{new} is the average accepted data traffic, and ρ_d is the average offered data traffic load (Erlangs). Clearly, $\overline{S^d}$ is actually the average throughput of the system. Similarly, $\overline{S^v}$ is defined where O_b^v depends on how many servers are busy with serving voice cells. Hence, by adapting the traffic load accordingly, the following quality of service parameters will be under control that are cell blocking probability call establishment delay and the cell error.

Now, it is straight forward to evaluate to the average cell delay, i.e.

$$\overline{D} = \frac{1}{\mu} \sum_{j_v} \sum_{j_d} \sum_{l=L+1}^{B} l \cdot X_l^d \cdot C(j_v, j_d) \tag{7.17}$$

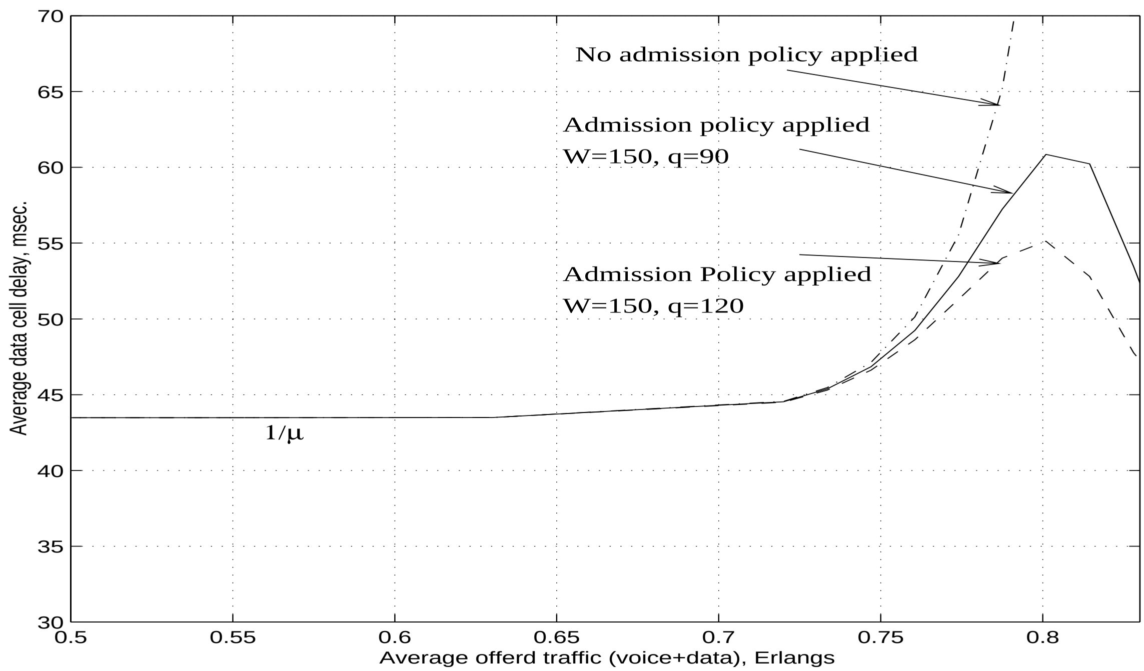


Figure 7.2: The average cell delay for different q window-based estimation.

where μ is the average service rate. Moreover, the average data cell error (P_e^d) and the voice cell loss (lv) is defined as follows.

$$P_e^d = \sum_{j_v} \sum_{j_d} (1 - u(j_v + j_d) \cdot d(j_v + j_d))\, C(j_v, j_d) \tag{7.18}$$

$$lv = \sum_{j_v} \sum_{j_d} \{1 - u(j_v + j_d) \cdot d(j_v + j_d) \cdot (1 - P(L))\}\, C(j_v, j_d) \tag{7.19}$$

where $P(L)$ is the probability of having all L servers busy.

7.4.3 Results and Discussion

Having considered the above system parameters, we first solve the queueing system of equations to obtain the steady state probabilities of buffer occupancy (i.e X_l^d under all possibilities of active voice and data calls. Then, using these probabilities, the performance measures are evaluated as mentioned above in equations (7.17-7.19), but without applying the adaptive admission policy. Then, the probability of

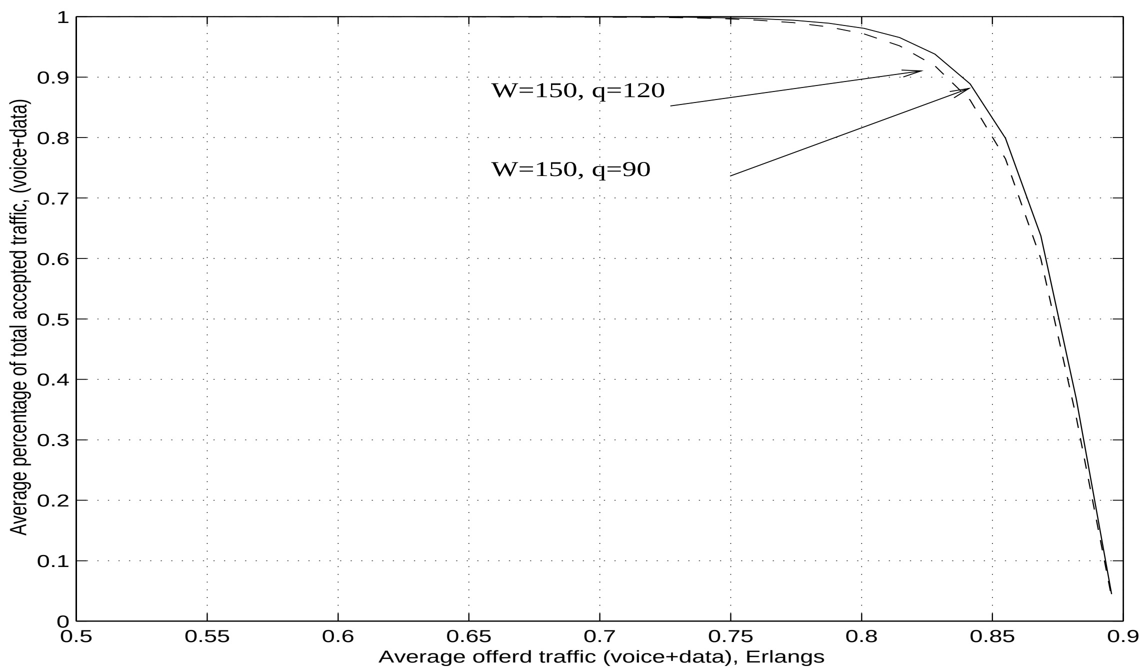


Figure 7.3: The average throughput for voice /data multiplexed traffic.

non-congested buffer is averaged over all possible population under certain arrival rates of voice and data. Then, the voice and data traffic will be adapted accordingly.

Figure 7.2 shows the average delay for different traffic load conditions ρ (i.e. $\rho = \theta_d \rho_d + \theta_v \rho_v$). It is clear our admission policy has controlled the average cell delay compared with the delay performance when no adaptive policy applied. Considering different windows requirements, we notice that the lower the number of consecutive cells q required to find the buffer in the state of non-congestion, the higher is the average cell delay. This is attributed to the fact that when q is low the call blocking is less and the consequently more cells will reside in the buffer. However, the price that should be paid for this good cell delay performance for data users is the low system throughput at high traffic loads as shown in Fig. 7.3.

Figure 7.4 compares the estimated buffer congestion under a window of W=150 cells and q= 90 cells. It is obvious that the proposed admission policy has effectively limited the cell blocking during the call. Finally, we examine the average data cell

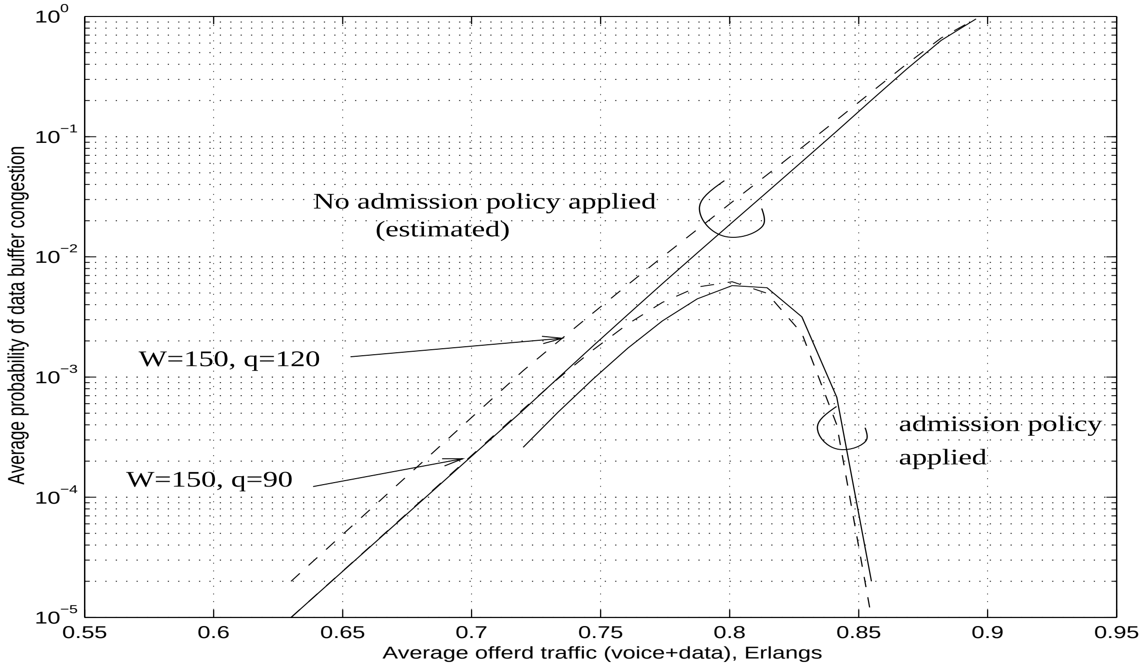


Figure 7.4: Comparison of the average data buffer congestion probability.

error rate and voice cell loss as shown in Figure 7.5, we notice that for all possible channel loads, the data cell error and voice loss have good performance and the proposed policy does maintain these critical QoS requirements. Although, we note that providing the base station with more servers on the down link to support high traffic load does not always help in maintaining the required QoS parameters. In fact, this is an interesting result and this happens because the more cells coexist on the down link, the higher is the probability of unsuccessful transmission.

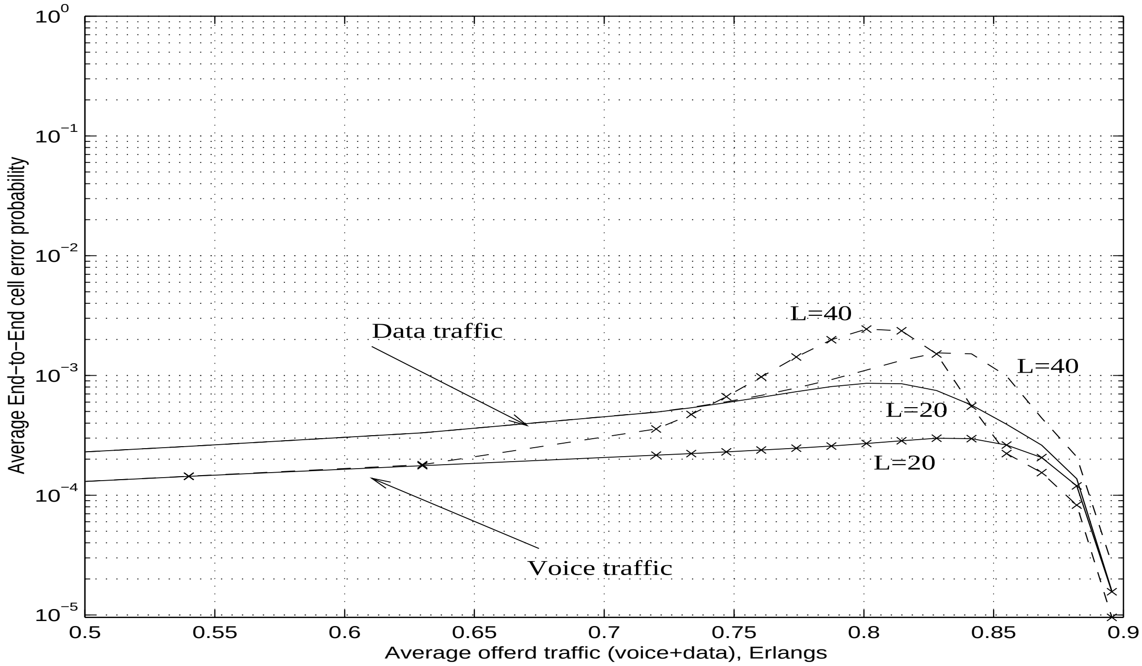


Figure 7.5: Comparison of the average End-to-End error cell probability, W= 150 cells, q= 90 cells, E_b/N_o = 7.5 dB.

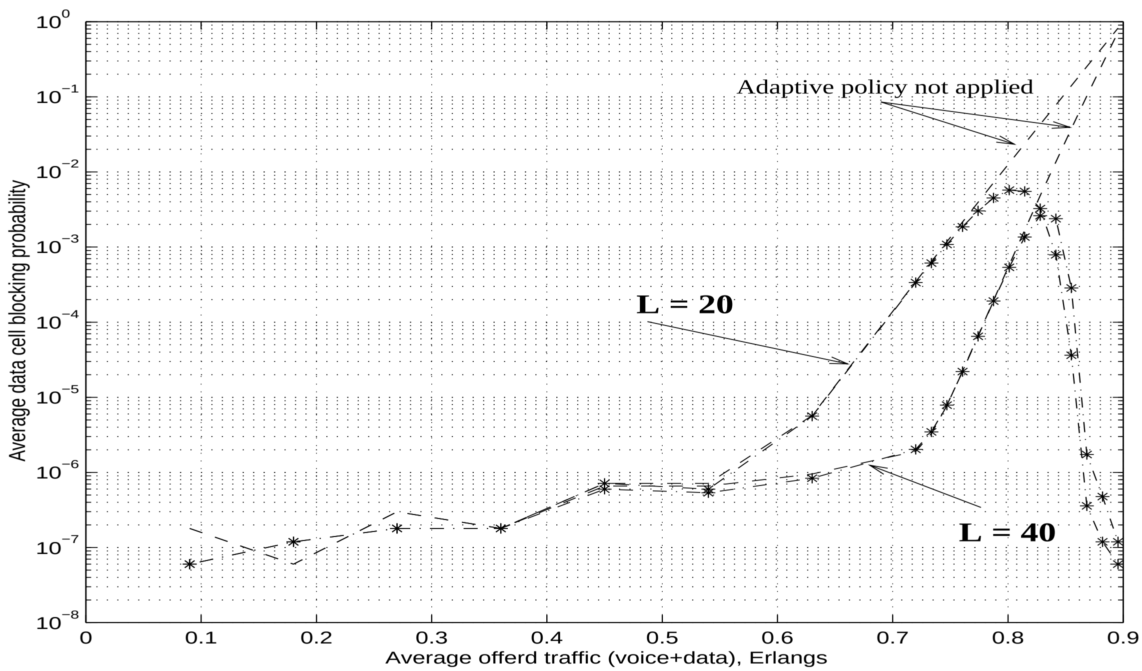


Figure 7.6: The average data cell blocking probability.

7.5 MC-CDMA/TDMA System Performance Analysis

In the following, we analyze the performance of the proposed TDMA/MC-CDMA system employing the 1st traffic flow control approach as explained in section 5.2.1. The analysis applies equally for both categories of traffic unless it is stated differently. Hence, and for analysis convenience, the system parameters are defined generally.

7.5.1 Performance Measures

The following MAC scenario is considered for evaluating the proposed admission / congestion policy. We assume that both the uplink and down link apply the two-slots CDMA/TDMA multiple access protocol. In addition, each slot population has its own buffer at the base station and each buffer shall be examined separately according to the window-based admission policy which will be presented shortly. This proposed scenario is illustrated in Fig. 7.7. It is important to note that mobile stations' receivers are conventional receivers (single-user detection). Therefore, the complexity involved in using multiuser detection is concentrated at the base station.

Following the analysis in chapter 6, it is then straightforward to evaluate the mean buffer occupancy for each slot population which takes into consideration all cells existing in the system (i.e. in service and in queue). We identify the average stream traffic buffer by '1' and the excess traffic buffer by '2'. Consider the average stream traffic buffer, we find $E_b^1(.)$ is given by

$$E_b^1(K_1, k_{11}, K_2, k_{22}) = \sum_{l=1}^{B} l X_l^1 \tag{7.20}$$

where X_l^1 is the solution of Eq. (6.12) using the algorithm for the average stream users where all cells are treated equally (no priority scheme applied), K_1 and K_2 are the total number of active users from both stream and interactive traffic admitted

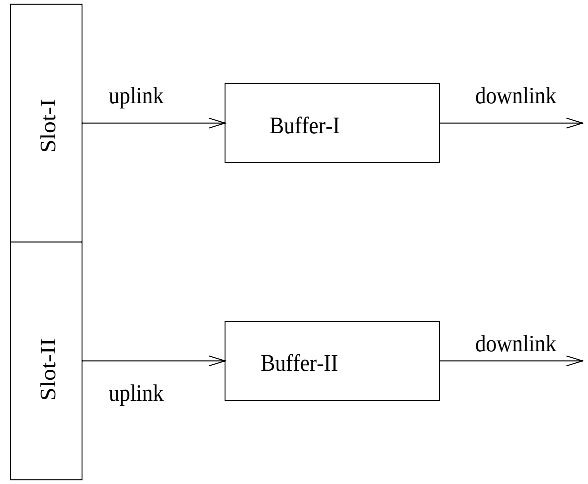


Figure 7.7: the proposed MAC scenario.

into the system, respectively. It is crucial to note that k_{11} and k_{22} are not single numbers, but they represent the indices for the sets of possible combinations of actual number of calls admitted from the stream traffic category and the interactive traffic category, respectively. For example, let $N^s = 4$ classes and there is one active stream user (i.e $K_1 = 1$), so $k_{11} \in \{(0,0,0,1),(0,0,1,0),(0,1,0,0),(1,0,0,0)\}$. As another example, let $N^{Int} = 2$ classes and there is two active interactive users (i.e $K_2 = 2$), so $k_{22} \in \{(1,1),(2,0),(0,2)\}$. Similarly, $E_b^2(.)$ is given by

$$E_b^2(K_1, k_{11}, K_2, k_{22}, j^{Int}) = \sum_{l=1}^{B} l X_l^2 \tag{7.21}$$

where j^{Int} is the instantaneous number of cells emitted from a specific combination of K_2 interactive users, and X_l^2 is the solution of Eq. (6.12) using the algorithm for the two-queue with higher priority is granted for interactive cells. From Eq. (7.21), we observe that the buffer content of the excess traffic is dependent on the instantaneous activities of admitted interactive users (i.e j^{Int}). Subsequently, $E_b^2(.)$ should be averaged over j^{Int} as shown in Eq. (7.22).

$$E_b^2(K_1, k_{11}, K_2, k_{22}) = \sum_{j^I=0}^{N^I} E_b^2(K_1, k_{11}, K_2, k_{22}, j^I)\lambda_j^{Int} \tag{7.22}$$

where λ_j^{Int} is the joint probability distribution of emitting j cells simultaneously from all active interactive users as defined in Eq. (7.8). Now we shall consider the average number of cells in the buffers waiting (queued) for service E_q^1 and E_q^2 for average stream traffic and excess traffic, respectively.

$$\begin{aligned} E_q^1(K_1, k_{11}, K_2, k_{22}) &= \textstyle\sum_{l=L+1}^{B} lX_l^1 \\ E_q^2(K_1, k_{11}, K_2, k_{22}, j^{Int}) &= \textstyle\sum_{l=L+1}^{B} lX_l^2 \end{aligned} \tag{7.23}$$

Subsequently, the waiting time in queue can be calculated as follows.

$$\begin{aligned} T_q^1(K_1, k_{11}, K_2, k_{22}) = \tfrac{1}{\mu} \textstyle\sum_{l=L+1}^{B} lX_l^1 \\ T_q^2(K_1, k_{11}, K_2, k_{22}, j^{Int}) = \tfrac{1}{\mu} \textstyle\sum_{l=L+1}^{B} lX_l^2 \end{aligned} \tag{7.24}$$

where μ is the service rate (cell/sec) which is deterministic.

Other important criteria in evaluating the overall system performance are the cell blocking probabilities $P_{bc}^1(.)$ and $P_{bc}^2(.)$, cell error performance $P_E^1(.)$ and $P_E^2(.)$ and the new call blocking probability P_{bn}^s. We define the cell blocking probability as the probability of finding the buffer full, i.e.

$$\begin{aligned} P_{bc}^1(K_1, k_{11}, K_2, k_{22}) &= X_B^1 \\ P_{bc}^2(K_1, k_{11}, K_2, k_{22}, j^{Int}) &= X_B^2 \end{aligned} \tag{7.25}$$

$$\begin{aligned} P_E^1(K_1, K_{11}, K_2, k_{22}) = (1 - u_m) + u_m(1 - d_m) \\ P_E^2(K_1, K_{11}, K_2, k_{22}, j^{Int}) = (1 - u_i) + u_i(1 - d_i) \end{aligned} \tag{7.26}$$

where m is the instantaneous number of cells from all K_1 users sharing slot-I (i.e. the average stream traffic), and i is the instantaneous number of cells from K_2 users including excess traffic from all K_1 users sharing slot-II. Further, we define Bl_1^s as the average new stream call blocking probability following phase-I of the admission policy, i.e.

$$P_{bn}^s(K_1, k_{11}, K_2, k_{22}) = 1 - \text{prob}(N^I + N_{excess}^s \leq L^u) \cdot \text{prob}(N_{avr}^s \leq L^u) \tag{7.27}$$

$$Bl_1^s = \sum_{K_1}\sum_{k_{11}}\sum_{K_2}\sum_{k_{22}} P_{bn}^s(k_1, k_{11}, k_2, k_{22}) \tag{7.28}$$

Similarly, all other performance measures are averaged over all possible combinations as will be shown later.

Regarding the interactive users, the only performance measures shall be considered here are the end-to-end cell (packet) error P_E^{Int} and the new call blocking probability P_{bn}^{Int}

$$\begin{aligned} P_E^{Int}(K_1, K_{11}, K_2, k_{22}, j^{excess}) &= (1 - u_i) + u_i(1 - d_i) \\ P_{bn}^{Int}(K_1, K_{11}, K_2, k_{22}) &= 1 - \text{prob}(N^I \leq L^u) \end{aligned} \tag{7.29}$$

where i is as defined above in Eq. (7.26) and j^{excess} is the instantaneous number of cells emitted from a specific combination of K_1 stream traffic and they are sharing slot-II along with interactive users.

The above mentioned performance measures indicate how the overall system works. Yet, there is no control scheme applied such that the QoS parameters are satisfied. Therefore, a congestion policy should be developed using one or more of the above performance measures.

Application of Window-Based congestion Policy

The policy presented in this subsection is applied equally for both buffers. Using the occupancy of the average stream traffic buffer (i.e. buffer 1), we define the cell overflow (or loss) probability $O_b^1(.)$, i.e.

$$O_b^1(.) = \sum_{l=TH^1}^{B} X_l^1 \tag{7.30}$$

where TH^1 is the ATM buffer content that defines congestion for buffer '1'. Equation (7.30) defines buffer overflow over one cell time. Similarly, we define the probability of non-congested buffer $S^1(.)$, assuming the dialup or call signaling period equals W cells and an incoming call finds the buffer in good condition for at least q consecutive cells, i.e.

$$S^1(.) \leq \sum_{i=0}^{W-q-2} \begin{pmatrix} W \\ i \end{pmatrix} (O_b^1)^i (1-O_b^1)^{W-i} + (W-q)^2 (O_b^1)^{W-q-1} (1-O_b^1)^{q+1}$$
$$+(W-q+1)(O_b^1)^{W-q}(1-O_b^1)^q \tag{7.31}$$

We note that $S^1(.)$ has the same parameters as the other performance measures. Then, $S^1(.)$ will be averaged over all possible combinations of admitted calls, i.e.

$$\overline{S^1} = \sum_{K_1} \sum_{k_{11}} \sum_{K_2} \sum_{k_{22}} S^1(K_1, k_{11}, K_2, k_{22})\ C(K_1, k_{11}, K_2, k_{22}) \tag{7.32}$$

where $C(.)$ is the joint probability distribution of a certain combination of active users initiated from both traffic categories. We repeat here the definition of the joint probability distribution of admitted calls of each class from all categories as defined in Eq. (7.10), assuming there are four independent classes of traffic belong to the stream traffic category, and two independent classes of traffic belong to the interactive traffic category.

$$C(K_1, k_{11}, K_2, k_{22}) = \prod_{k=1}^{4} Q_k^s(i_k) \cdot \prod_{k=1}^{2} Q_k^{Int}(j_k) \tag{7.33}$$

where $K_1 = \sum i_k$ corresponds to the total number of calls from the stream traffic users and $(K_2 = \sum j_k)$ corresponds to the total number of calls from the interactive traffic users. $Q_k^s(i_k)$ and $Q_k^{Int}(j_k)$ are the probability distributions for the stream and interactive traffic categories, respectively, as defined in Eq. (7.11). For the purpose of simplicity in calculating the average of $S^1(.)$ for all possible combinations which are very huge, we make the following approximation. We assume that for each admitted pair of K_1 and K_2 users, all their possible combinations are equally likely, i.e.

$$S^1(K_1, K_2) = \frac{1}{k_{22}k_{11}} \left(\sum_{k_{22}} \sum_{k_{11}} S^1(K_1, k_{11}, K_2, k_{22}) \right) \tag{7.34}$$

Then, $S^1(K_1, K_2)$ shall be averaged over all possible K_1 and K_2. As it was mentioned before, in this study, the probability distribution of cell generation as well as the call initiation are assumed to be Bernoulli distribution with different activity factors (i.e. θ, δ, respectively).

$$\overline{S^1} = \sum_{K_1}\sum_{K_2} S^1(K_1, K_2)C(K_1, K_2) \tag{7.35}$$

$$C(K_1, K_2) = \Gamma^s_{K_1} \cdot \Gamma^{Int}_{K_2} = \sum_{i_1=0}^{Su_1^s}\sum_{i_2=0}^{S_2^s}\sum_{i_3=0}^{Su_3^s}\sum_{i_4=0}^{Su_4^s}\prod_{k=1}^{4} Q_k^s(i_k) \cdot \sum_{j_1=0}^{Su_1^{int}}\sum_{j_2=0}^{Su_2^{Int}}\prod_{k=1}^{2} Q_k^{Int}(j_k) \tag{7.36}$$

where Su_1^s and Su_1^{Int} are the numbers of subscribers to class 1 in the stream traffic category and interactive traffic category, respectively, and so on. This approximation will be used in all calculations of all other performance measures. We should note that the above analysis is equally applicable for buffer '2'.

Since S^1 and S^2 refers to the same traffic category (i.e. stream category) but evaluated from two different buffers, we follow a conservative approach where the lower one will be used to control the flow of traffic (i.e. $\overline{S^s} = \min(\overline{S^1}, \overline{S^2})$).Therefore, the traffic load must be adapted according to $\overline{S^s}$ as shown in (7.37).

$$\rho^s_{new} = \rho^s \cdot \overline{S^s} \tag{7.37}$$

Where ρ^s_{new} is the average accepted traffic, and ρ^s is the total stream offered traffic load (Erlangs).

$$\rho^s = \frac{1}{L}\sum_{i=1}^{N^s} \delta_i^s \cdot \xi_i^s \cdot N_i^s \cdot \theta_i^s \tag{7.38}$$

where L is the maximum number of downlink servers. Similarly, the total interactive offered traffic load (Erlangs) ρ^I is defined, i.e.

$$\rho^I = \frac{1}{L}\sum_{i=1}^{N^{Int}} \delta_i^{Int} \cdot \xi_i^{Int} \cdot N_i^{Int} \cdot \theta_i^{Int} \tag{7.39}$$

Now we are in a position to state our new call admission policy; i.e., the new admission states "The new call is admitted once the total number of low-bit streams corresponding to the active calls is less than M as in equation (7.28) and the buffer is not congested as in equation (B.5)." The overall new call blocking probability in this case is given by

$$Bl_2^s = P^s_{block} = 1 - \overline{S^s} \tag{7.40}$$

Table 7.1: Traffic Characteristics and System Parameters

	Stream Traffic # of Classes= 4	Interactive Traffic # of Classes= 2
Class 1 (cell level)	ξ_1^s= 3, $\hat{\xi}_1^s$= 2 N_1^s= 20, $\theta_{1,on}^s$= 1.0	ξ_1^I= 7 N_1^{Int}= 10, $\theta_{1,on}^{Int}$= 0.5
Class 2 (cell level)	ξ_2^s= 20 kbps, $\hat{\xi}_2^s$= 2 N_2^s= 25, $\theta_{2,on}^s$= 0.75	ξ_2^I= 50 kbps N_2^{Int}= 10, $\theta_{1,on}^{Int}$= 0.35
Class 3 (cell level)	ξ_3^s= 4, $\hat{\xi}_3^s$= 3 N_3^s= 10, $\theta_{3,on}^s$= 0.90	
Class 4 (cell level)	ξ_4^s= 5, $\hat{\xi}_4^s$ = 4 N_4^s= 5, $\theta_{4,on}^s$= 0.80	
Bandwidth, BW	5 MHz	5 MHz
Basic rate, R_b	10 kbps	10 kbps
Spreading code	gold code	gold code
Processing gain, PG	255	255
Packet size	53 bytes	53 bytes
Maximum # of downlink servers (L)	60	60
Service rate μ cells/sec	47	47

Therefore, the second phase of the proposed policy (congestion) has two steps. First, the likelihood of buffer congestion ($\overline{S^s}$) is estimated over a period of W cell-time unites. Second, the offered traffic load is adapted accordingly. Similar to the argument presented in section 7.2.1, the mean call establishment time, i.e.

$$\begin{aligned}\overline{CE^s} &= \textstyle\sum_{l=1}^{\infty} l \cdot (P_{block}^s)^{l-1}(1 - P_{block}^s) \cdot (W + \tau) \\ &= \tfrac{(W+\tau)}{(1-P_{block}^s)} = \tfrac{(W+\tau)}{\overline{S^s}} \quad \text{cell times}\end{aligned} \tag{7.41}$$

where τ is the round trip propagation delay between user and BS (or satellite), involved in each user signaling trial of duration W cells.

7.6 Results and Discussion

In evaluating the performance of the aforementioned admission/congestion policy on MC-CDMA/TDMA access protocol, we consider the following system specifications. We assume that the uplink traffic and the downlink traffic are using different channels. The source stream bits are modulated using BPSK-DS-CDMA with a processing gain of 255 using different Gold codes. The low-bit average stream traffic bits are transmitted via the uplink without encoding (multiuser decorrelator receivers are used at the BS) while these bits shall be encoded at the BS by convolutional codes for transmission on the down link channel as will be explained later. More, it is assumed that these average stream packets transmitted via the uplink will only be distorted by AWGN ($E_b/N_o = 12$ dB) besides the mutual interferers from other intracell users.

On the other hand, the excess traffic bits as well as the interactive traffic bits are encoded by convolutional codes. Further, following the **IS-95** specifications [99], the uplink bits ($E_b/N_o = 8.2$ dB) are encoded by a convolutional code where $k = 9, R = \frac{1}{3}$, while the downlink bits ($E_b/N_o = 8.45$ dB)are encoded by another convolutional code with $k = 9, R = \frac{1}{2}$. It is assumed that the cell on the uplink channel are received noncoherently at the base station. On the other hand, the transmitted cells via the downlink channel are assumed to be received coherently by the mobile stations. Both links are assumed to be under the influence of two paths Rayleigh multipath fading. Table 7.1 summarizes the traffic characteristics and system parameters used in evaluating the proposed traffic control approaches.

Further, we assume that the number of system subscribers to both traffic categories is limited as shown in table 7.1. In our results, we let the basic rate $R_b = 10$ kbps to be the same for all system users. Though, in the case of two slots CDMA/TDMA and because of the 2 slots framing, the bursty rate should be higher than R_b. Here, we assume that the bursty rate at each time slot is $2R_b$ which can be translated into approximately 47 ATM cells/sec. It is also assumed that no new call

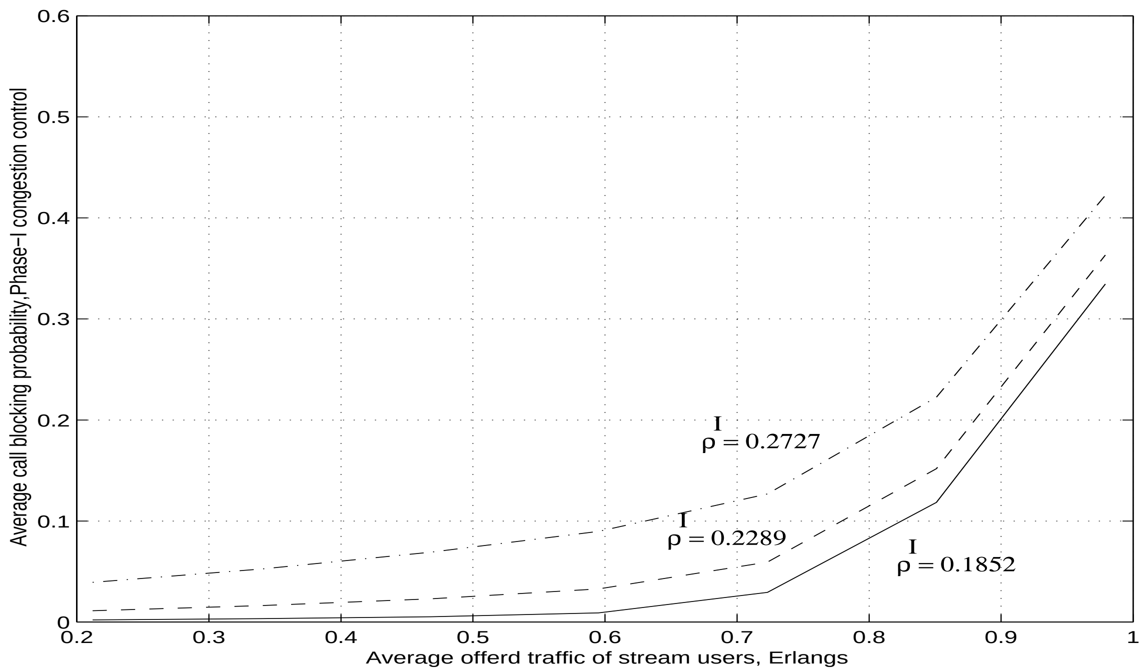


Figure 7.8: The average call blocking probability of stream traffic under the phase-I of congestion control policy.

is initiated by the busy user until the current call is completed and the user becomes idle again. In this work, interactive users calls have higher priority over the stream traffic users calls. In the case of contention, the interactive cells will be served first if there are available servers, otherwise, these cells will be rejected. Therefore, the interactive cells suffers no delay waiting in the queue. Each buffer has a capacity (i.e. B) of 70 cells. However, the thresholds for buffer overflow are different where $TH^1 = 61$ (i.e. $TH^1 = L + 1$) and $TH^2 = \frac{2}{3}B$. $TH^1 = L + 1$ is chosen to be just greater than the maximum number of available servers. On the other hand, we chose lower threshold for buffer '2' which its contents come from excess traffic such that it gives an early warning of the build up process of cells in the buffer such the performance of interactive users do not degrade severely and at the same time, the overall stream traffic users have acceptable performance. The measurement window W equals to 150 cells ($\approx$ 6.36 sec.), and q is taken to be 90 cells.

Having considered the above system parameters, we first solve the queueing

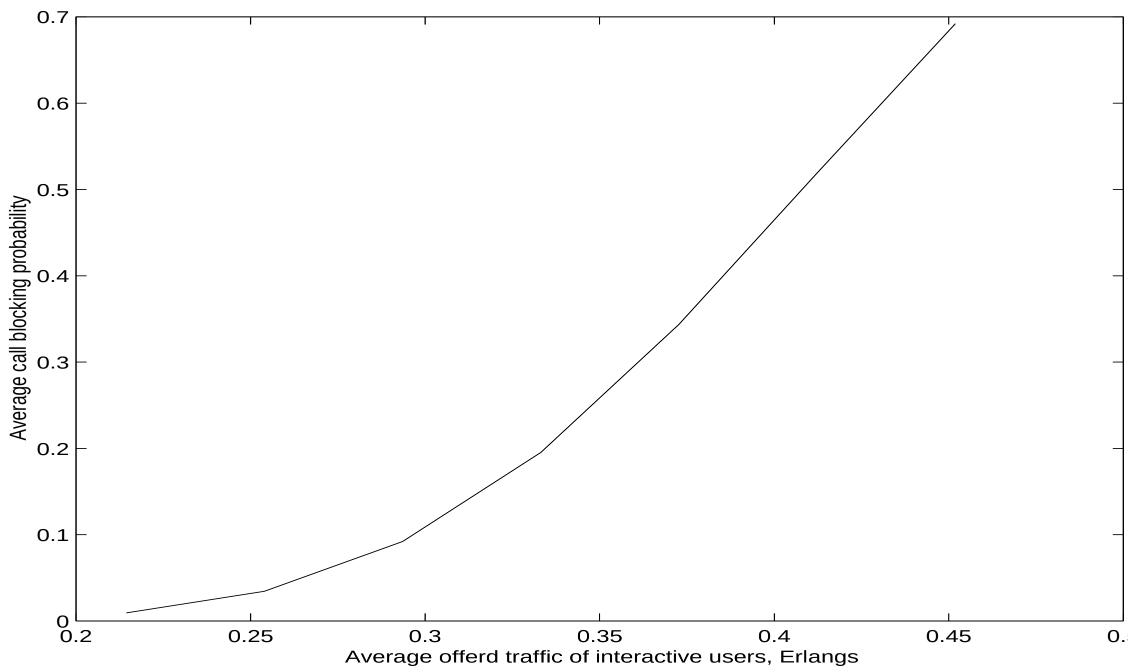


Figure 7.9: The average call blocking probability of interactive traffic under the phase-I of congestion control policy.

system of equations to obtain the steady state probabilities under all possibilities of active calls. Then, using these probabilities, the probability of non-congested buffer as defined in Eq. (B.5) is evaluated for both buffers (i.e. $\overline{S^1}$ and $\overline{S^2}$). Now, the stream offered traffic load should be adapted according to the estimated probability of congested buffer $\overline{S}$ (i.e. $\rho^s_{new} = \rho^s \cdot \overline{S}$). Subsequently, the performance measures are evaluated as mentioned above in equations (7.21-7.29) for possible combinations of admitted pair of (K_1, K_2) active users. Then, these measures shall be averaged as we have done in equations (7.32-7.36).

Figures 7.8 & 7.9 show the call blocking probability under Phase-I (admission) policy for both stream and interactive traffic categories, respectively. The higher the interactive traffic load, the higher is the call blocking probability for the stream traffic users calls. Figure 7.10 illustrates the probability of buffer congestion estimated by the window-based admission/congestion policy. It shows that the average system throughput for stream traffic load deteriorates for high stream traffic load. Also, it

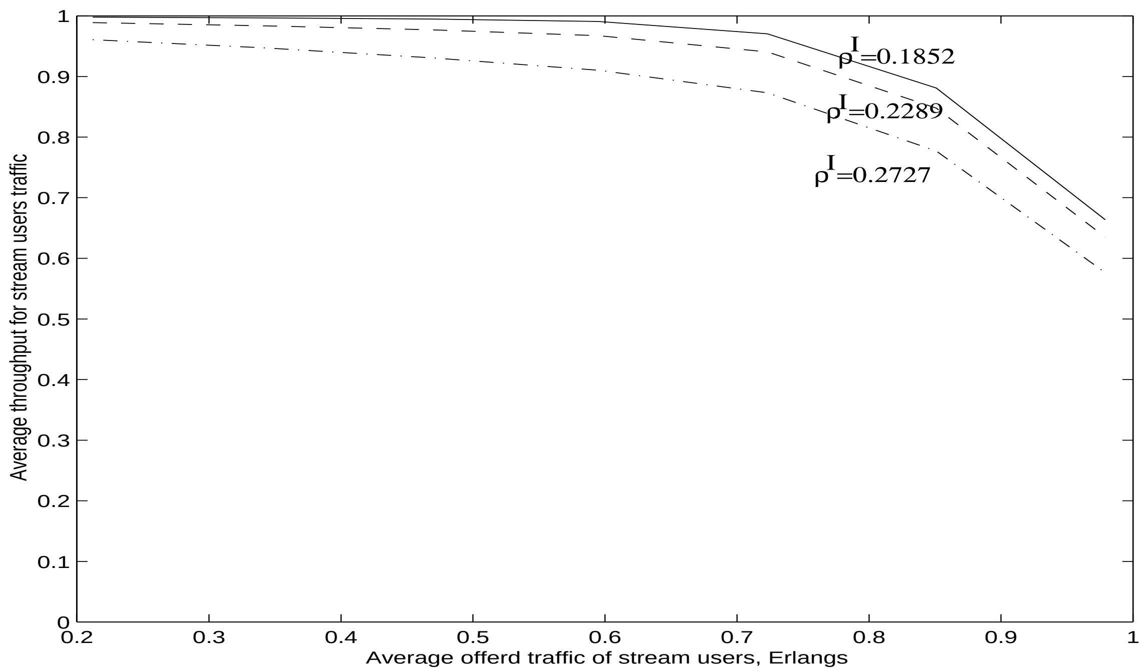


Figure 7.10: The average throughput of stream traffic under the admission/congestion control policy

points out to the influence of the interactive users on the overall system performance. The higher the interactive traffic load, the lower are the accepted calls. Therefore, we are trading off the system throughput for maintaining other QoS parameters such as cell delay, call establishment delay, cell error, etc. Figure 7.11 shows the average buffer occupancy of both slots' buffers. It is clear that the average stream buffer contents is much higher than the other and this is due to nature of the flow traffic control applied for slot-I traffic. However, at high traffic load, the buffer occupancy is under control. Consequently, the average cell delay is within reasonable range as shown in Fig. 7.12. We note that Fig. 7.12 shows the average cell delay due to the queued cells (i.e. $\overline{E^s_q}$) and not the total delay ($\overline{D}_t = T_s + \frac{1}{2}T_F + \overline{T^s}_q$) experienced by the transmitted cells via slot-I, where T_s is slot duration in seconds, and T_F is the frame duration in seconds (in our study $T_F = 2T_s$). Of course, the first two terms are constant and the population of both slots (i.e. slot-I and slot-II) always experience that portion of the delay. It is an important to explore the relation between the average cell delay and the achieved throughput. This relation is shown in Fig. 7.13. It is obvious that for low and moderate traffic loads, both performance measures (i.e. delay and throughput) are good, but for high traffic load the throughput should be sacrificed such that an acceptable cell delay be maintained. Moreover, the delay improvement at high loads ρ^s shown in Fig. 7.10 are mainly attributed to the combined call blocking and the congestion control policy (i.e. on the expense of more call blocking as shown in Fig. 7.8). Further, Fig. 7.14 shows the average call establishment delay. Again, it is obvious that the proposed policy has controlled the call establishment delay and maintained it in the range of one window for all expected traffic loads.

Now we examine the cell error performance for both traffic categories. Figures 7.15-7.17 show the average cell error performance for the cells transmitted via slot-I, the variable-rate components (excess traffic) of the stream users transmitted via slot-II and the cell error performance for interactive users traffic, respectively. It is interesting to note that the proposed flow traffic control along with the new admission/congestion policy effectively maintains the cell error for both traffic categories within acceptable ranges. For instance, cells transmitted via slot-II suffer average cell error less than 0.005 for all possible traffic loads. On the other hand, cells transmitted via slot-I (average stream traffic) have good cell error performance except at high traffic load. This can be attributed to the traffic flow control policy where during slot-I, the mobile station persists on transmitting a fixed number of cells (average) such that the multiuser decorrelator receiver implementation becomes more practical. Nevertheless, it is obvious that the admission/congestion policy works well in maintaining the cell error performance within the acceptable ranges, but with sacrificing the overall system throughput. Finally, Fig. 7.18 shows the average cell blocking probability when the traffic is adapted according to the $\overline{S^s}$. On the other hand, the average cell blocking for excess traffic was found to be approximately zero. This happens due to the low excess traffic population.

We can easily recognize from the results presented here that due to the nature of the proposed system (MC-CDMA/TDMA, traffic flow control, and admission/congestion policy) both traffic categories have influence on each other. At high interactive load, less stream traffic shall be admitted which results in lower cell error, lower cell blocking, less delay, and so on. However, the price that is paid is lower throughput.

We have noticed that (not presented herein) providing the base station with more servers on the down link (while keeping the same spread spectrum bandwidth) to support high traffic load does not always help in reducing the cell delay. In fact, this is an interesting result and takes place because the higher is the number of cells on the down link, the higher is the probability of unsuccessful transmission.

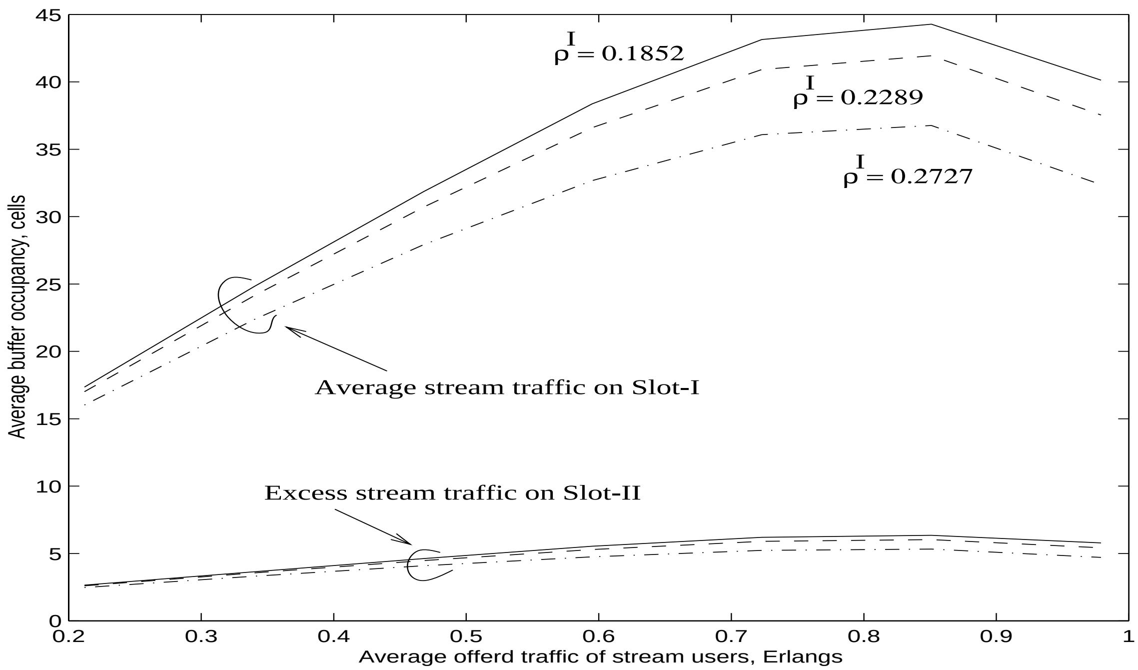


Figure 7.11: The average buffer occupancy of stream traffic users under the admission/congestion control policy.

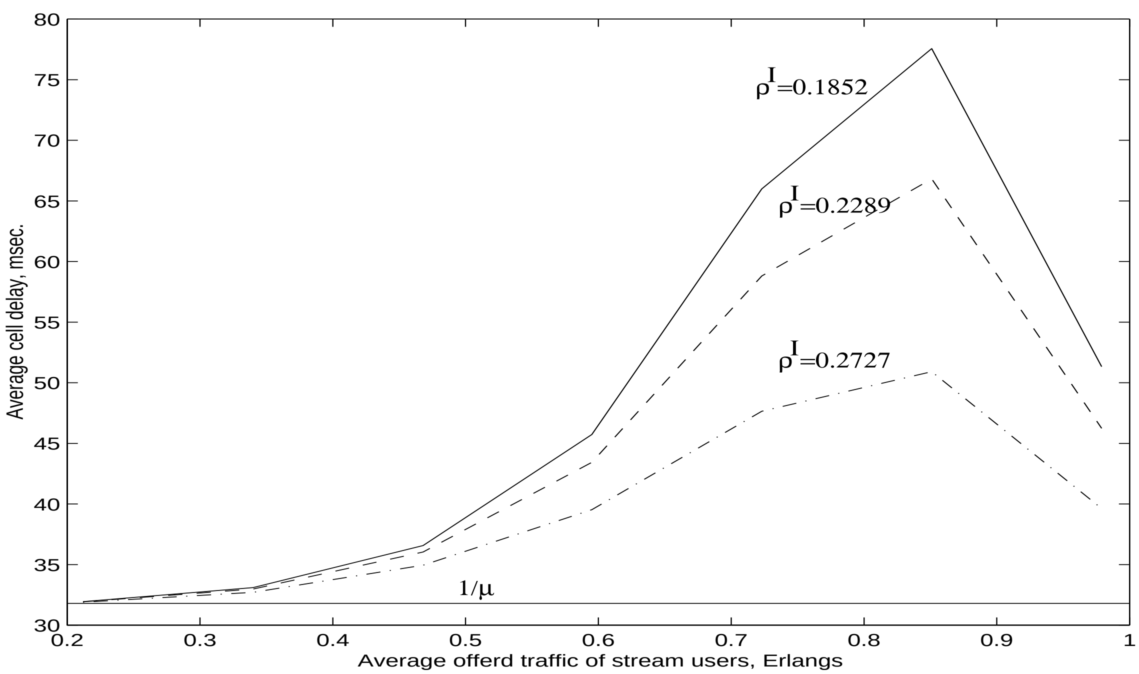


Figure 7.12: The average cell delay of stream traffic under the admission/congestion control policy

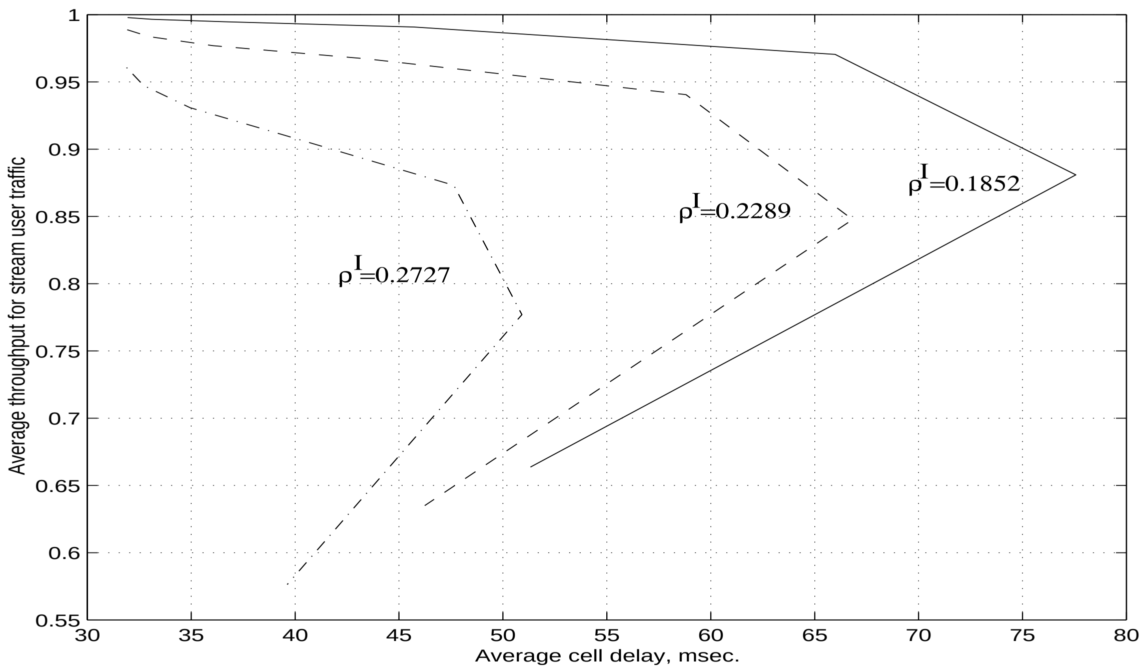


Figure 7.13: The average throughput versus the average delay of stream traffic under the admission/congestion control policy

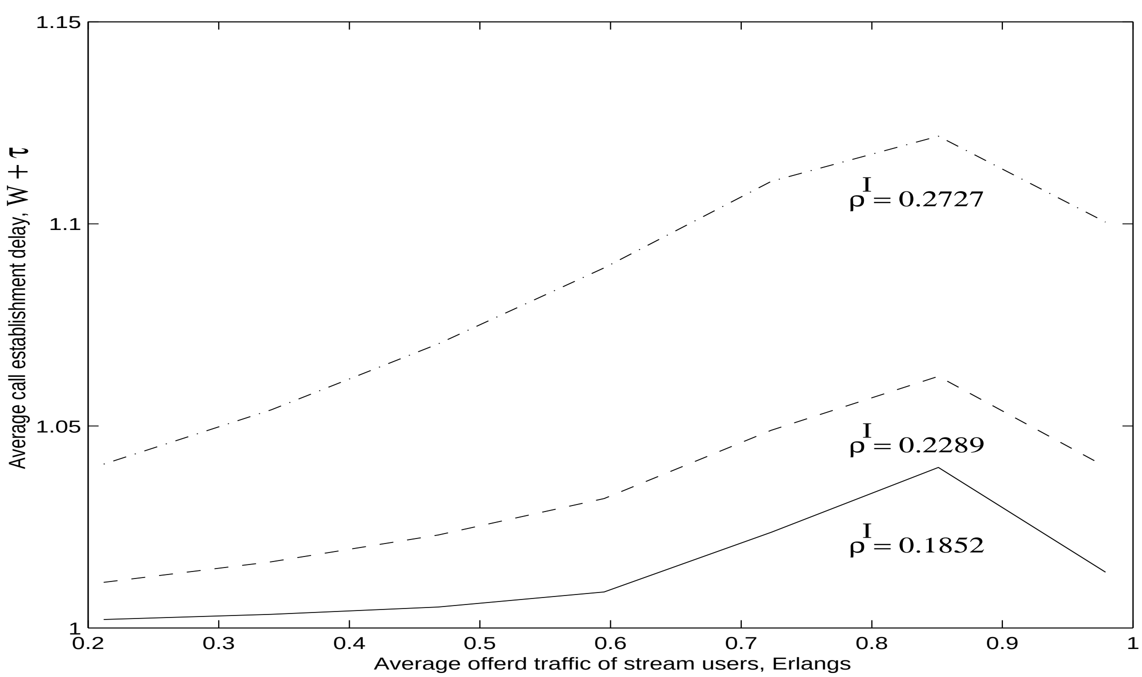


Figure 7.14: The average call establishment delay for stream traffic users.

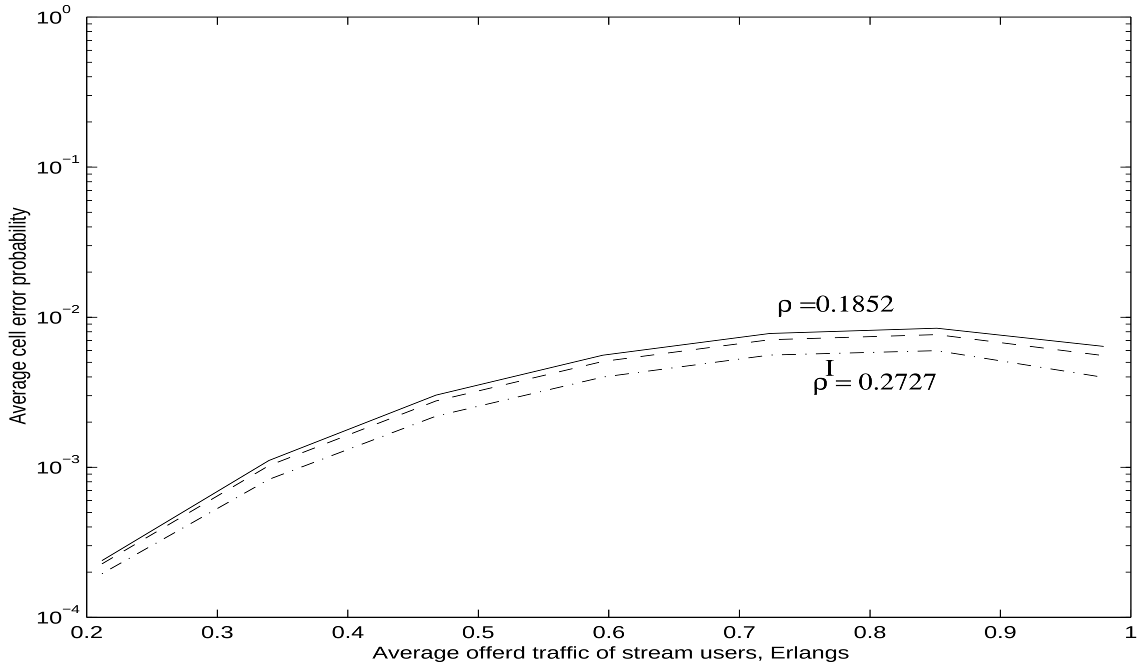


Figure 7.15: The average cell error probability of average stream traffic transmitted via slot-I.

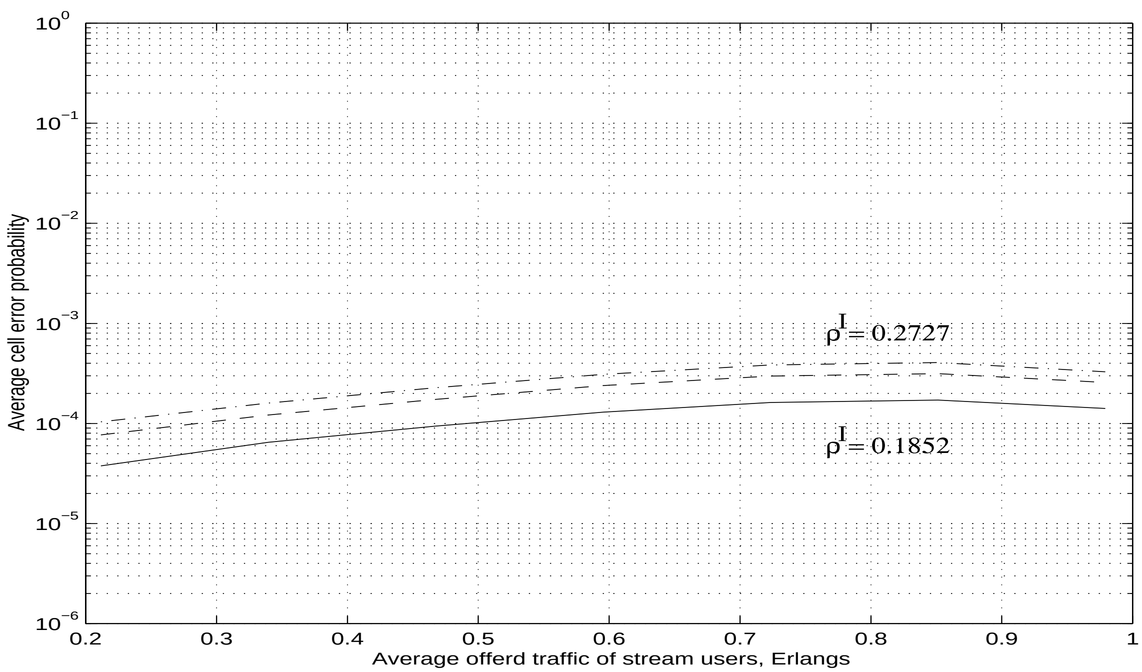


Figure 7.16: The average cell error probability of excess stream traffic transmitted via slot-II along with active interactive cells.

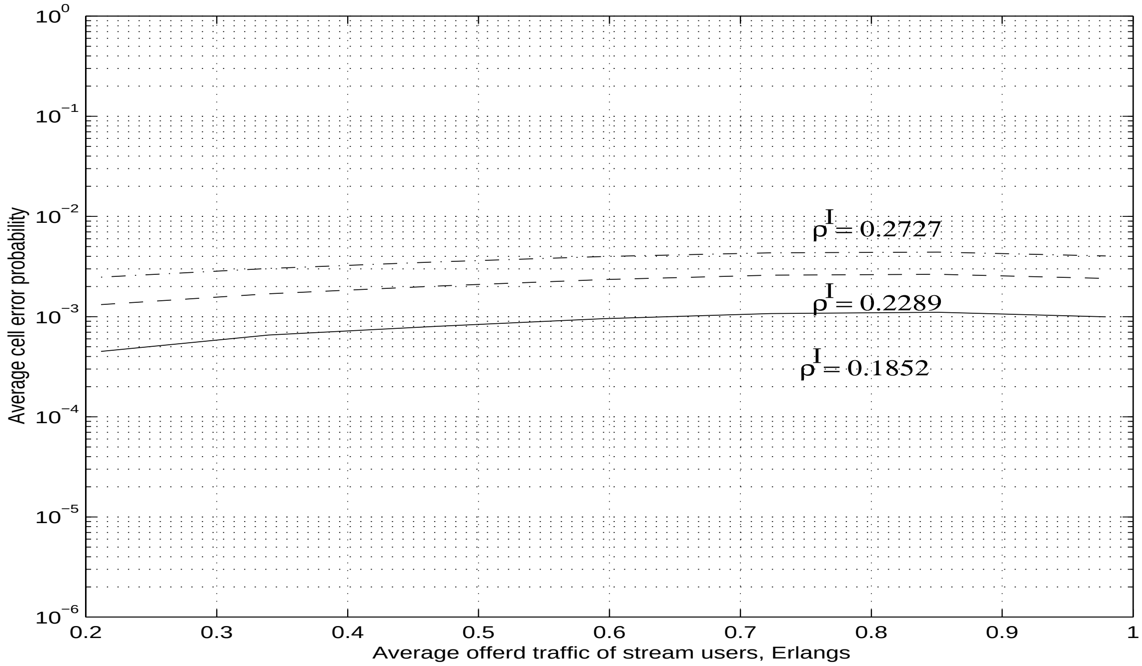


Figure 7.17: The average cell error probability of interactive traffic transmitted via slot-II along with excess stream traffic cells.

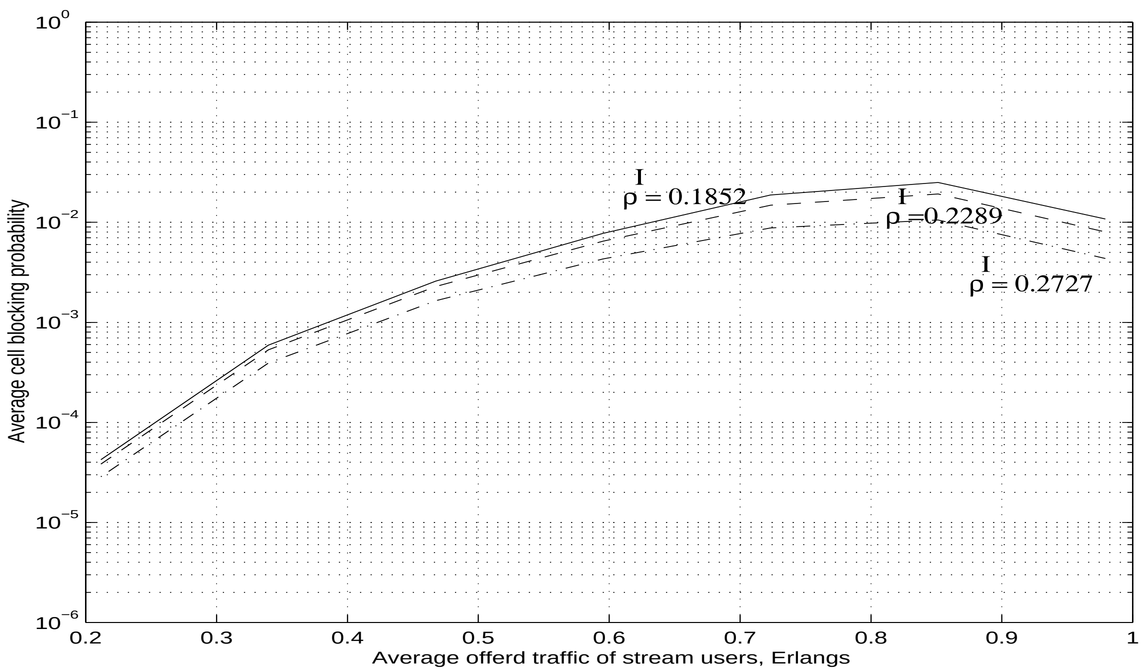


Figure 7.18: The average cell blocking probability of the average stream traffic buffer (i.e. Slot-I).

Chapter 8

Simulation Study for the Proposed MC-CDMA/TDMA

8.1 Introduction

In the previous chapters, we have analyzed analytically the proposed access protocols accompanied with window-based measurement admission/congestion control policy. The complexity of the wireless networks under consideration has put some restrictions on our analysis. Hence, it is important to evaluate the performance of the proposed protocols via simulation where these protocols shall be tested under more realistic conditions and system parameters.

8.2 System Description

We use exact Monte-Carlo simulation to simulate the multimedia integrated CDMA networks where heterogeneous traffic users are multiplexed into a simple TDMA frames. Each frame is composed of two slots as explained before. In this simulation, our emphasis is on the network performance issues such as packet delay, packet delay jitter, packet losses, call blocking, etc. Therefore, the only system parameters that are going to be simulated and randomly generated throughout the simulation program are these parameters which have direct relation to the network performance. For instance, packet error performance will not be simulated. However, the Monte-Carlo simulation is supplemented by analytical computation of packet error probability for uplink stream packets received by decorrelator receiver at the base station (BS). On the other hand, the packet error probability for the uplink slot-II users as well as downlink users from both slots are supplemented from computational analysis given by [99] which we briefly summarize in Appendix A.

In Fig. 8.1, we depict a general system block diagram for the CDMA network under consideration. We show only the main blocks related to our work. Since our main focus is concentrated on the network issues, no details will be given regarding source coding nor channel coding. Indeed, the implications of these blocks are implicitly considered in the packet error performance. Further, it is worth to emphasis that we concentrate on the bulk of packets transmitted via uplink channel by all admitted mobile stations and the same thing applies for the down link channel.

Assume a set of $\mathcal{A}$ users are admitted to the network. The process of packet generation is a random process which will be explored later. At every packet-time unit, there will be a group of packets $\mathcal{B}$. As shown in Fig. 8.1, this flow of information bits shall be spread by signature codes $[C]$ and then they will be modulated using **BPSK** modulation technique which formulates the S_u signal, i.e.

$$S_u = diag(\mathcal{B}\,\mathcal{C}) \cdot \mathcal{F}_u^c \tag{8.1}$$

where

$$\begin{aligned} \mathcal{B} &= [b_1(t)\ b_2(t) \cdots]^T \\ \mathcal{C} &= [c_1(t)\ c_2(t) \cdots] \\ \mathcal{F}_u^c &= [\cos(\omega_c t + \phi_1) \quad \cos(\omega_c t + \phi_2) \cdots] \end{aligned} \tag{8.2}$$

S_u shall experience the impairment of the uplink channel (fading, multipath fading, shadowing, etc.). At the BS, a distorted version of S_u is received (i.e. R_u). It was stated before that the proposed admission policy interrelates the end-to-end system parameters such as packet error rate on uplink channel as well as downlink channel. Hence, we are concerned about how much of those transmitted packets via uplink channel (i.e. $\mathcal{B}$) could be received successfully at the BS (i.e. $\hat{\mathcal{B}}$). Of course $\hat{\mathcal{B}} \in \mathcal{B}$.

For each simultaneous transmitted packets $\mathcal{B}$, we read the corresponding packet error rate P_e as explained in Appendix A. Then a uniform random generator (RG) is triggered to determine whether these packets could arrive successfully at the BS. For instance, let $P_e = 10^{-5}$, the output of the uniform RG shall fall between 0 and 1. We consider the event '0' (failed transmission) if the RG output is $\leq P_e$ and then these packets will be discarded and not considered as an input to the BS buffer (i.e. $\hat{\mathcal{B}} = 0$). Otherwise, it is considered '1' (successful transmission) and consequently these successful packets shall be included in the BS buffer input (i.e. $\hat{\mathcal{B}} = \mathcal{B}$).

At the BS, $\hat{\mathcal{B}}$ packets will be treated differently depending on the admission policy applied as well as type of received packets (for example, voice or data).

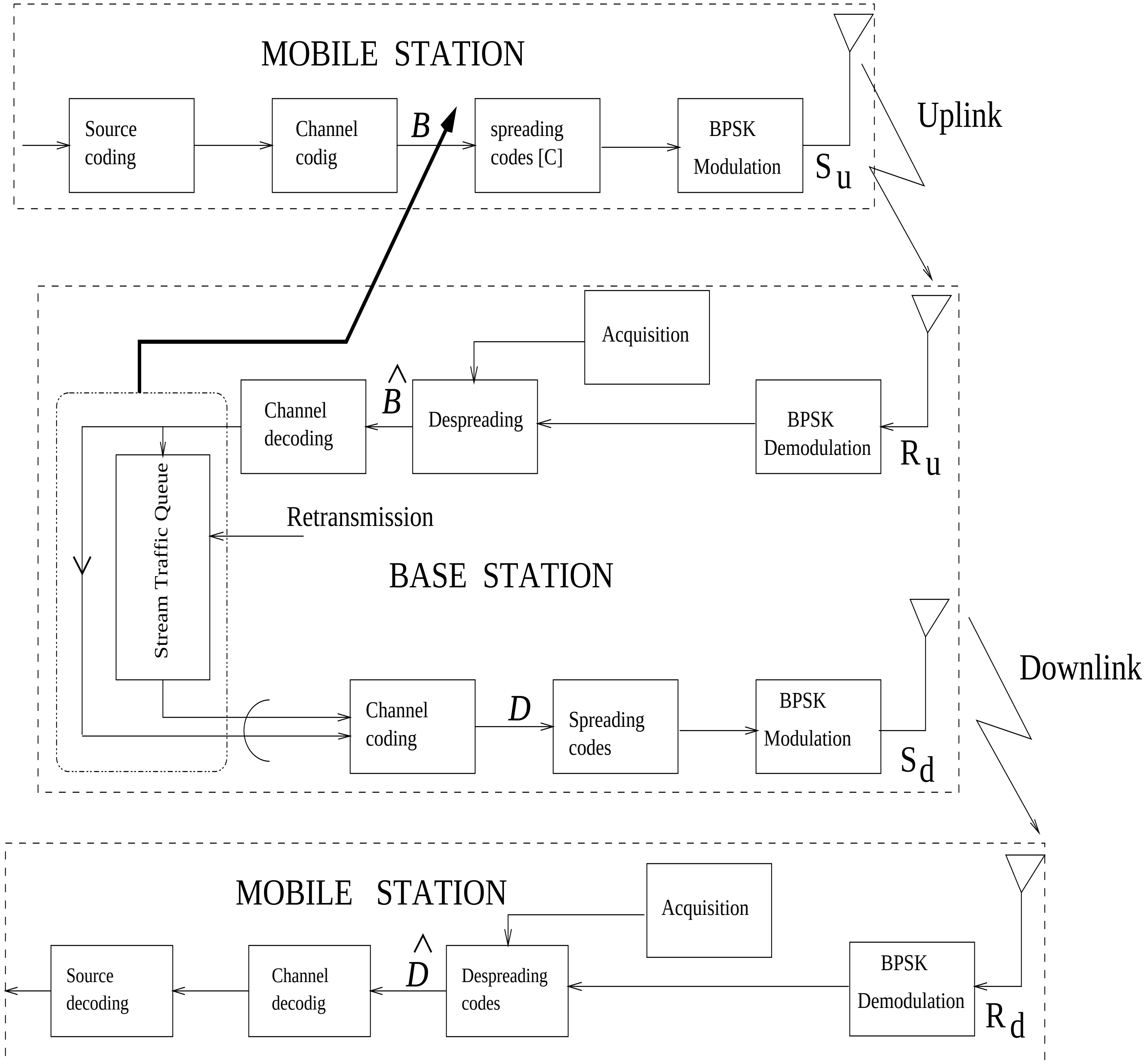


Figure 8.1: System block diagram

Considering packets received during slot-II, $\hat{\mathcal{B}}$ is composed of two flows of packets: 1) packets belong to *excess* traffic and 2) packets belong to *interactive* traffic. The later will be served immediately if there are available servers on the down link. Otherwise, these packets should be dropped and they are not allowed to be queued. However, the former (excess packets) will be queued in two cases. First, if there is no available servers on the down link provided that there is enough space in the buffer. Otherwise, these packets should be dropped also. Second, if the transmission on the down link was not successful, copies of those lost packets will be queued and retransmitted later. Figure 8.1 denote $\mathcal{D}$ as a vector containing streams of packets from excess as well as interactive traffic who are permitted for transmission via the downlink channel. Then, $\mathcal{D}$ packets will be processed in a similar way of $\mathcal{B}$ and results S_d.

$$S_d = diag(\mathcal{D}\ \mathcal{C}) \cdot \mathcal{F}_d^c \tag{8.3}$$

where $\mathcal{D}, \mathcal{C}, \mathcal{F}_d^c$ are defined similarly to those in Eq. (8.1). Here, we assume that the uplink channel and downlink channels are using different frequency bands. Similarly, S_d will suffer the impairment of the down link channel and will be distorted and received as R_d by the other end-mobile.

8.3 Simulation Assumptions

Up to this point, it should be clear that our emphasis is on '**packet level**'. In other words, the simulation concentrates on the following random events: 1) Call initiation, 2) packet generation (i.e. $\mathcal{B}$), 3) successfully transmitted packets via uplink channel $\hat{\mathcal{B}}$, 4) BS operation (i.e. packet dropping, buffer overflow, busy servers on downlink channel) $\mathcal{D}$, and 5) successfully transmitted packets via downlink channel $\hat{\mathcal{D}}$.

8.3.1 Simulation of Calls initiation

A new call request is generated by a uniform random generator with certain activity probability. Since our observations are carried out every window, the activity probability is the average call activity over a window (for example, 0.005 calls/window). The aggregated calls generated from all independent users yield a multinomial random process. We assume the call duration is geometrically distributed. So, a *geometric* RG is triggered to find the call duration for this specific call in terms of *measurement windows*. Here, we impose a simple modification for the sake of shortening the time required for each simulation run that is the calls initiated at the same window from the same traffic class shall have the same call duration. Moreover, these calls associated with their duration are registered in a special register and it shall be updated by the end of each window.

8.3.2 Simulation of Packets Generation

Throughout our analysis in the previous chapters, traffic are assumed to be heterogeneous and each source is modeled as a single ON/OFF source for the sake of computational simplicity. Yet, this assumption does not take into consideration the correlation between generated packets. Hence, in this simulation, our traffic sources of interest are assumed to have variable bit rates. Each source has its own activity (i.e α, β) and by which its bit variations is controlled. Nevertheless, a constant bit rate user still can be modeled as *variable bit rate* source with $\alpha = 0$.

Fluid Source Modeling of Traffic Sources [100]

The *Fluid Source Modeling* is chosen for simulating heterogeneous sources under consideration in this work for two main reasons. The first reason is that the source itself could be a *variable bit rate* source. Hence, the simultaneous number of emitted packets is varying. The second reason is our *multicode* transmission policy, even

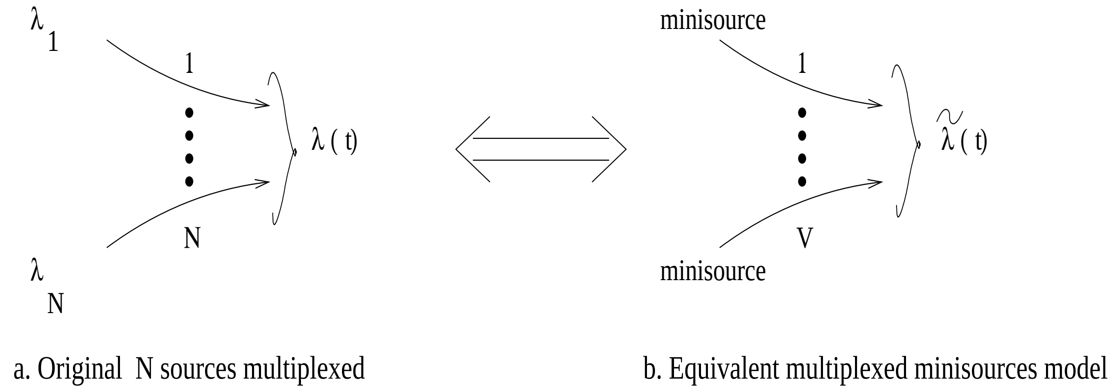


Figure 8.2: Equivalent process

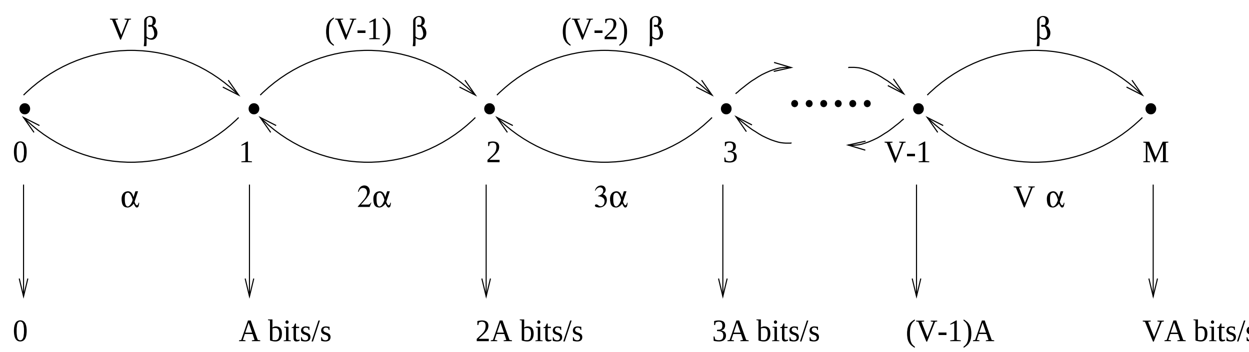


Figure 8.3: Markov chain representation of the equivalent process

though, the source is a *constant bit rate*.

Each source is represented by an *equivalent* process. The equivalent process will be defined to be of V identical two-state “mini-sources”, each moving back and forth exponentially between an “off state” and an “on state” in which A bits/s are generated. In other words, the time-varying bit rate of each source is quantized into V quantization levels. To reduce the effect of quantization, the number of mini-sources should be increased (i.e $V >> N$). In this simulation, we set A to be equal to the basic transmission rate (i.e R_b). Figures 8.3.2 and 8.3.2 illustrate the Markov chain representation of the equivalent process.

The simulated kth received signal is

$$S_{k,j}^{M}(t) = \sum_{i=1}^{\xi_j} \left(c_{k,j}^{i}(t) b_{k,j}^{i}(t) cos(\omega_c(t) + \phi_k) \right) \otimes h_{k,j}(t) \tag{8.4}$$

This is the compound signals coming from the kth user belonging to the j calls using the Mth slot for transmission; $M = 1, 2$. Further, ξ_j is R_j/R_b, and needless to mention that ξ_j is limited by the practical application (for example, the existing spreading codes, etc.), ϕ_k is the phase shift associated with the kth user. We note that ξ_j is unit-less.

In evaluating the performance of the aforementioned admission/congestion policy, we consider the following system specifications. We assume that the uplink traffic and the downlink traffic are using different channels. The source stream bits are modulated using BPSK-DS-CDMA with a processing gain of 255 using different Gold codes. The low-bit average stream traffic bits are transmitted via the uplink without encoding (multiuser decorrelator receivers are used at the BS) while these bits shall be encoded at the BS by convolutional codes for transmission on the down link channel as will be explained later. More, it is assumed that these average stream packets transmitted via the uplink will only be distorted by AWGN ($E_b/N_o = 12$ dB) besides the mutual interferers from other intra cell users.

On the other hand, the excess traffic bits as well as the interactive traffic bits are encoded by convolutional codes. Further, following the **IS-95** specifications [99], the uplink bits ($E_b/N_o = 8.2$ dB) are encoded by a convolutional code where $k = 9, r = \frac{1}{3}$, while the downlink bits ($E_b/N_o = 8.45$ dB) are encoded by another convolutional code with $k = 9, r = \frac{1}{2}$. It is assumed that the packet on the uplink channel are received non-coherently at the BS. On the other hand, the transmitted packets via the downlink channel are assumed to be received coherently by the mobile stations. Both links are assumed to be under the influence of two paths Rayleigh multipath fading. Table 8.1 summarizes the traffic characteristics and system parameters used in evaluating the proposed traffic control approaches.

Table 8.1: Traffic Characteristics and System Parameters

	Stream Traffic # of Classes= 4	Interactive Traffic # of Classes= 2
Average call length (geometric)	1000 packets	3 min
Class 1 (packet level)	ξ_1^s= 10, $\hat{\xi}_1^s$= 7 N_1^s= 100, $\theta_{1,on}^s$= 0.7	ξ_1^I= 10 N_1^{Int}= 10, $\theta_{1,on}^{Int}$= 0.25
Class 2 (packet level)	ξ_2^s= 4, $\hat{\xi}_2^s$= 3 N_2^s= 100, $\theta_{2,on}^s$= 0.75	ξ_2^I= 5 N_2^{Int}= 10, $\theta_{1,on}^{Int}$= 0.36
Class 3 (packet level)	ξ_3^s= 8, $\hat{\xi}_3^s$= 7 N_3^s= 100, $\theta_{3,on}^s$= 0.90	
Class 4 (packet level)	ξ_4^s= 6, $\hat{\xi}_4^s$ = 4 N_4^s= 100, $\theta_{4,on}^s$= 0.80	
Basic rate, R_b	10 kbps	10 kbps
Spreading code	gold code	gold code
Processing gain, PG	255	255
Packet size	53 bytes	53 bytes
Maximum # of servers (L)	60	60
Service rate μ packets/sec	47	47

8.4 Simulation Block Diagram

8.4.1 Course of Simulation

At any time, initiated call requests are either accepted or rejected. The waiting users shall be served first if it is possible according to the buffer status. Then, we check if there is any new call requests that could be served. If the buffer status shows that the probability of buffer congestion is high, we have the following scenarios. First, we keep the priority for those users who are waiting (of course, for a certain 'Time out' period) and block any new call requests. The second scenario is to add the new calls to the waiting list.

The simulation procedure for the call admission can be divided into two parts. One part checks the new call requests pool, while the other checks whether there are any standing calls from previous windows. Figures 8.4.1and 8.4.1 depict the block diagram of Phase-I of the course of simulation. These diagrams are applicable for both time slots unless it is stated otherwise. Here, we shall explain the functionality of each block, its parameters and its influence on the network performance.

Each simulation run is composed of K equal **Measurement Windows**. At the beginning of each window, new users can initiate admission requests to establish communication with the network. As soon as a user generate a call request and could not be admitted solely because of buffer condition, his Call Establishment Delay (CED) counter is initialized and it shall be incremented through the simulation program until the call is accepted. The final value of this counter is an indication of the number of packets that the concerned user had to wait until its call is accepted So, the average call establishment delay $CED(j)$ for all system users during the jth simulation run is

$$CED(j) = \frac{\text{total CED per user * \# of active users}}{\text{total accepted calls}} \tag{8.5}$$

Then, the BS should check the capacity needed by each new user and compare it to the availability of servers on the uplink as well as the buffer condition at the BS.

Here, we should elaborate on the mechanism used in this simulation model to check the capacity needed by the user. As stated in several places in this work, the users from both traffic categories could have *variable bit rate* (VBR) or *constant bit rate* (CBR) users. However, each user has its own transmission characteristics. Each user is modeled as mini-sources ON/OFF model. The number of the mini-sources is different from user to user. Hence, the flow of information rate from each user has a maximum (R_{max}) and minimum (R_{min}) rates. In this policy, we compare the total rates of incoming users plus the total rates of existing users with the available number of servers. Here, we may have different possibilities of defining the rate. For instance, we can say

$$R = R_{min} + Const * (R_{max} - R_{min})/R_{min} \tag{8.6}$$

for different *Const*. For now, we are using the peak rate of each user. Notice that this policy has to be followed in both slots.

For example, assume there are three new call requests belonging to the first class of stream traffic, and there are other two call requests from the first class of interactive traffic. More, assume that queueing is not possible at the user side. Hence, it is essential to accommodate these users even at their maximum rates. Now, we have to calculate three parameters, namely, the maximum possible packets emitted from interactive users, average packets and maximum possible excess packets from stream users. Referring to table 8.1, we find the following: Av_stream = $3\xi_1^s = 21$, Ex_stream = $3(\xi_1^s - \hat{\xi}_1^s) = 9$, Intr = $2\xi_1^I = 20$. Therefore, the expected traffic from these new users for slot-I is 21 packets while for slot-II is 29 packets. Now, assume that the demands of previously admitted calls are as follows: 20 servers on slot-I and 40 servers on slot-II. Now, if the stream traffic has higher priority than the interactive traffic, then, all new stream traffic can be accepted directly, while new interactive shall be rejected.

The second phase in the admission policy is to examine the buffer status. If it is found that the buffer is in a 'good' condition, then these new calls can be admitted to the network. Here, we define the "good" condition as the state when the buffer occupancy is less than the maximum number of servers on the downlink. Once the call is accepted, the following should be done: the ON-call register is updated, where all accepted calls are registered. Subsequently, this information is fed back into the other slot where its users can control their traffic accordingly. In contrary, if the first condition (i.e. availability of servers) is not satisfied, the call will be blocked and the call-blocked register is updated (i.e. $C_b = C_b + 1$) and then the simulation program proceeds to evaluate the system performance under the current traffic load.

On the other hand, if the first condition is guaranteed but the buffer is not in a 'good' condition (the second criterion), we suggest the following policy. This caller shall be registered in a temporary record (i.e. Reg-temp) such that he is granted higher priority (during the next window) than later callers and a '*Time-wait*' counter

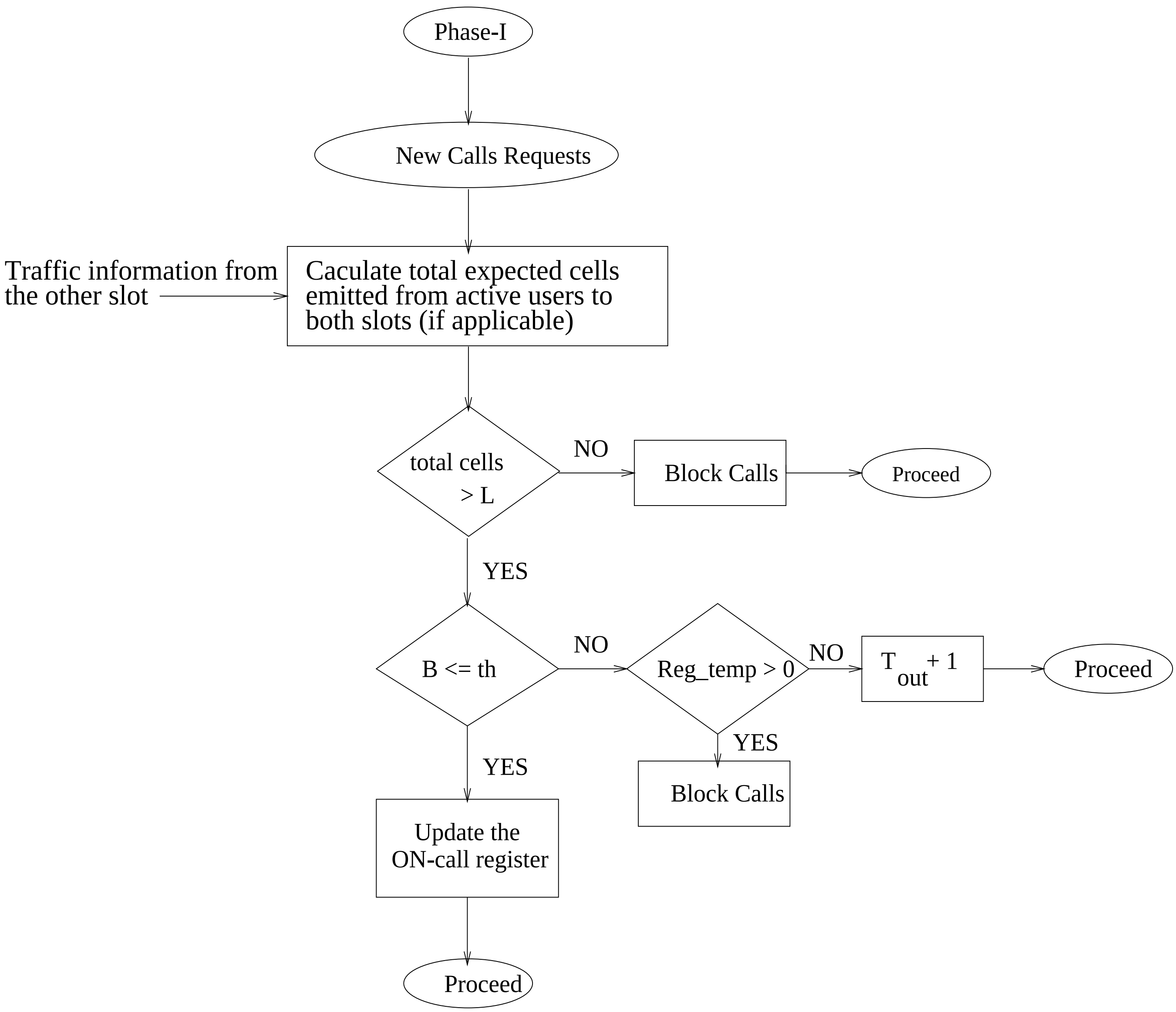


Figure 8.4: The simulation block diagram, Phase-I (Part-I)

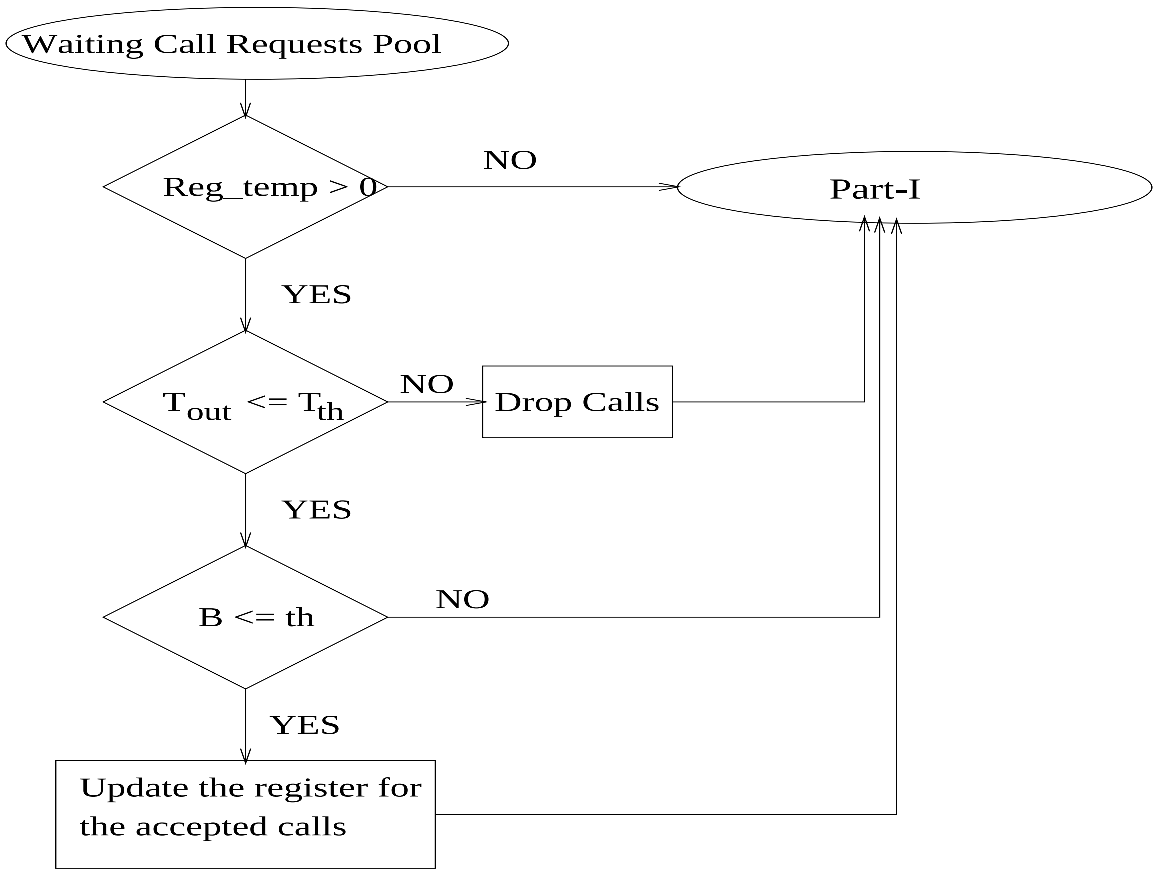


Figure 8.5: The simulation block diagram, Phase-I (Part-II)

is initiated to control the period in which this user is granted higher priority. The motivation behind this policy is to trade the call blocking probability ($C_b(j)$) and the call establishment delay ($CED(j)$). Further, by accepting a longer delay to establish a call we could reduce C_b probability by allowing the active users to statistically be multiplexed. Nevertheless, if the '*Time-wait*' exceeds a certain limit '*Time-out*', the following actions should take place: the call is dropped and this caller is returned into the potential users pool, the call-blocked register is updated (i.e $C_b = C_b + 1$), and the temporary register (i.e. Reg-temp) is updated.

Referring to the above example, accepting these new calls depends on the buffer condition as well as the availability of servers on the uplink. Let us assume that there are already active calls whose needs are as follows: 24 servers on slot-I and 30 servers on slot-II. Therefore, the maximum (peak) number of packets that would be transmitted is 45 (i.e. $24 + 21$) and 59 in slot-I and slot-II, respectively. Then, it obvious that as far as the buffer is in a 'good' condition, these new calls should be accepted. Nevertheless, assume that the buffer is not in a 'good' condition, then, all stream traffic requests should be registered in a temporary register where they shall wait for certain period until they are accepted or rejected, while interactive calls can be accepted straightforwardly. Of course, there are other possible scenarios.

Now, it is clear that the call-block counter (i.e. C_b will be incremented under two circumstances. Firstly, if this call can not be severed due to a shortage in servers on the uplink. Failing to get access to the network (after a time-out period) due to a buffer congestion is the second cause of call blocking. Further, as it will be explained in section 8.4.3, C_b counter will also be updated under another circumstance which is due to network congestion control policy. Under this policy the network operator may choose to terminate the new calls after one measurement window because otherwise the quality of service of currently carried calls would be degraded. Therefore, at the end of the simulation time (ST), the $C_b(j)$ is given by

$$C_b(j) = \frac{C_b}{\text{total call requests}} \tag{8.7}$$

Then the simulation proceeds to evaluate the system performance. We note, however, that in Fig. 8.4.1, we add a block to check if there are any calls waiting from previous windows and then if there is, the new calls should be blocked. This is a simple modification to the above policy such that the *CED* is put under control.

Now, we turn to the other half of the simulation block diagram of Phase-I, which begins with checking the *Waiting Call Requests* where a group of users who have been delayed to get access to the network from previous windows due to the current status of the buffer are waiting. We check first if those users do not violate '*Time-out*' threshold. If it is so this group of calls should be blocked and their callers return to the potential callers pool. Then, the simulation program progresses as explained above. However, if these waiting calls do not violate the '*Time-out*' threshold and the buffer condition is 'good', then these calls will be admitted to the networks. Further, the simulation program checks if there is any new call that can be accommodated as explained in the first part of the simulation block diagram. On the other hand, if the buffer is still not in a 'good' condition, the '*Time-wait'* counter is updated and the simulation proceeds to check if there is any new call requests.

It is important to measure how many packets the user would loss during the course of his call. The impairment of the uplink and downlink channels as well as buffer congestion at the BS are the main causes for packet loss. In this simulation, we have defined the following loss probabilities. First, the probability of packet loss due to unsuccessful transmission via uplink channel during the jth simulation run (i.e. $l_{up}(j)$) is given by

$$l_{up}(j) = 1 - \frac{\text{total packets received successfully at the BS}}{\text{total packets transmitted via uplink channel}} \tag{8.8}$$

Second, the probability of packets loss due to unsuccessful transmission via downlink channel during the jth simulation run (i.e. $l_d(j)$) is given by

$$l_d(j) = 1 - \frac{\text{total packets received successfully by the other mobile user}}{\text{total packets transmitted via downlink channel}} \tag{8.9}$$

Third, the packet losses due to buffer congestion during the the jth simulation run (i.e. $cell_b(j)$) which has a crucial influence on the network performance and is given by

$$cell_b(j) = \frac{\text{packets forced to drop because the BS buffer overflow}}{\text{total packets received successfully at the BS}} \tag{8.10}$$

From the user point view, it is very vital to measure the average overall packet losses. In other words, we should find the end-to-end packet received successfully, which may be given by

$$Packet_{throughput}(j) = \frac{\text{total number of end-to-end successfully received packets}}{\text{total generated packets}} \tag{8.11}$$

Moreover, it is important from the network design point of view to measure the $throughput(j)$ of the network for the jth simulation run which can be defined as follows.

$$throughput(j) = \frac{\text{total accepted traffic}}{\text{average offered traffic}} \tag{8.12}$$

where total accepted traffic is the actual number of packets generated by all accepted users. On the other hand, the offered traffic is the expected number of packets which would be generated by all accepted users if their calls had not been terminated due to the application of the admission/congestion policy. Since the simulation time is limited (only 60 min), this factor is taken into consideration in finding the average offered traffic where some calls might be terminated not because of the application of admission/congestion policy but because the simulation run ends. Finally, due to the impairment of the down link channel and the priority granted for interactive users, some stream packets might reside in the buffer for a certain period of time. Hence, these buffered packets should be counted and we can easily find the average packet delay $D(j)$ for the jth simulation rum, i.e.

$$D(j) = \frac{\text{number of packets waiting in the buffer}}{\text{number of packets permitted to the BS buffer}} * \text{frame time} \tag{8.13}$$

More, This performance measure could be also be used to control the flow of traffic to assure that the delay requirements by stream users are not violated.

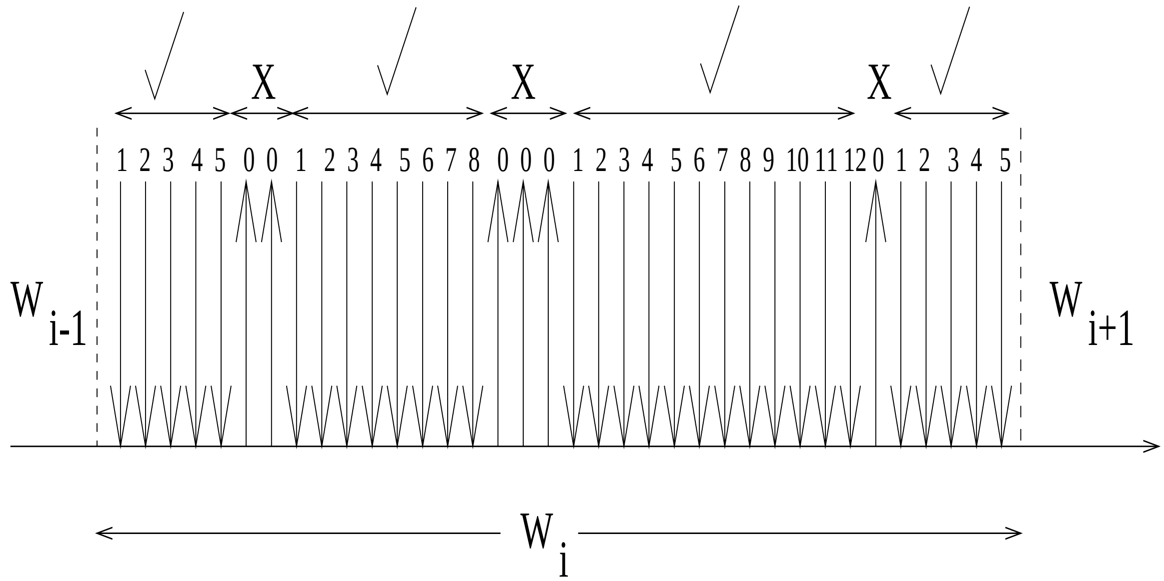


Figure 8.6: Illustration of buffer activities at the BS; window size= 36 packets.

8.4.2 Admission/Congestion Policies

Figure 8.6 shows an ensemble of network activity at the base station. We denote the case when the number of packets in the buffer exceeds certain threshold th_B by an upward arrow. Otherwise, it is denoted by a downward arrow and this what we call the buffer is in 'good' condition. This randomness is due to many factors, such as the traffic burstiness at the packet level, the erroneous channels, etc. These activities also vary from window to window due to the above reasons as well as the call activity of each user.

Therefore, it is obvious that these activities could give us a valuable information about how the whole network is working. Consequently, traffic loads might be controlled in such a way that the QoS requirements will be met. Hence, three *Admission/Congestion policies* are proposed.

Firstly, the admission of new users is dependent on the buffer condition at the end of each *measurement window* regardless of what was going on throughout the window. Of course, this policy is relax and weak. we call this policy the basic admission/congestion control policy.

In our analytical analysis, we have adopted a stringent admission policy that requires the buffer to stay in the 'good' condition for q consecutive *packet-time* units. For instance, applying this policy for W_i in Fig. 8.6, we find that frequency of occurrence of this event is

$$S = \frac{\text{\# of events}}{\text{Maximum \# of possible events}} \tag{8.14}$$

where the maximum number of possible events for a window of size W and q consecutive packets is given by

$$\Psi(W,q) = \sum_{i=0}^{W-q+1} \lfloor \frac{W-i}{q} \rfloor \tag{8.15}$$

where $\lfloor \cdot \rfloor$ is the greatest integer. Then, considering the example at hand (illustrated in Fig. 8.6), we can observe several small windows where the buffer is not congested for at least q consecutive cells (i.e packets). The first small window has a size of 5 packet-time units and the buffer were during this small period not congested. For this example, this small window corresponds to two events and so on.

$$S = \frac{2\sum_{i=0}^{2} \lfloor \frac{5-i}{4} \rfloor + \sum_{i=0}^{5} \lfloor \frac{8-i}{4} \rfloor + \sum_{i=0}^{12} \lfloor \frac{12-i}{4} \rfloor}{\sum_{i=0}^{33} \lfloor \frac{36-i}{4} \rfloor} = \frac{16}{33} = 0.48 \tag{8.16}$$

This above policy can be modified to obtain a moderate admission policy whose performance lies between the above policies. The modification is as follows. We just count the events where the buffer occupancy exceeds certain threshold. Then, we find how frequent does this event happen throughout the whole *measurement window*, i.e.

$$S = 1 - \frac{\text{\# of event occurrences}}{\text{Measurement window size}} \tag{8.17}$$

From Fig. 8.6, we find for the example above, that $S = 1 - (6/36) = 0.833$.

8.4.3 Traffic Adaptation

Let ρ_o^i be the total accepted traffic in window i (W_i), i.e.

$$\rho_o^i = \rho_a^{i-1} + \rho_{new}^i \tag{8.18}$$

where ρ_a^{i-1} is the accepted offered traffic arriving during W_{i-1}, and ρ_{new}^i is the new offered traffic during W_i. According to the original proposed admission policy (see section 7.2), the accepted traffic will be a ratio of the offered traffic as follows,

$$\rho_a^{i+1} = S \cdot (\rho_a^{i-1} + \rho_{new}^i) \tag{8.19}$$

Here, we modify the original admission policy such that a priority discipline is included. The old accepted calls shall be given higher priority than the new traffic and this is because of the fact that these excess packets belong to a call that is already active in the other slot. So, it is preferable to block a new call rather than terminating an active one. Therefore, it is easy to modify the admission policy accordingly and get the following relation,

$$\rho_a^{i+1} = S^{'} \cdot \rho_{new}^i + S^{''} \cdot \rho_a^{i-1} \tag{8.20}$$

We set $S^{''} = 1$ and by solving equations (8.19) and (8.20), we obtain the relation between the modified S and the traffic intensity from each traffic category, i.e.

$$S^{'} = (1 + \nu) \cdot S - \nu \tag{8.21}$$

where ν is the ratio of the old accepted load to the new traffic load (i.e. $\nu = \rho_a^{i-1}/\rho_{new}^i$). Figure 8.7 shows the relation between the original and modified admission policies. We note from Fig. 8.7 that there is a threshold for the values of $S^{'}$ if it is exceeded, $S^{'}$ is not valid any more. It is very easy to find this threshold in terms of ν,

$$S \geq \frac{\nu}{1 + \nu} \tag{8.22}$$

If the above condition is not satisfied, then all new traffic shall be blocked (i.e. $S^{'} = 0$) as well as a portion of the old accepted calls. Hence, we should solve equation (8.20) for $S^{''}$, i.e.

$$\rho_a^{i+1} = S^{''} \cdot \rho_a^{i-1} = S \cdot (\rho_a^{i-1} + \rho_{new}^i) \tag{8.23}$$

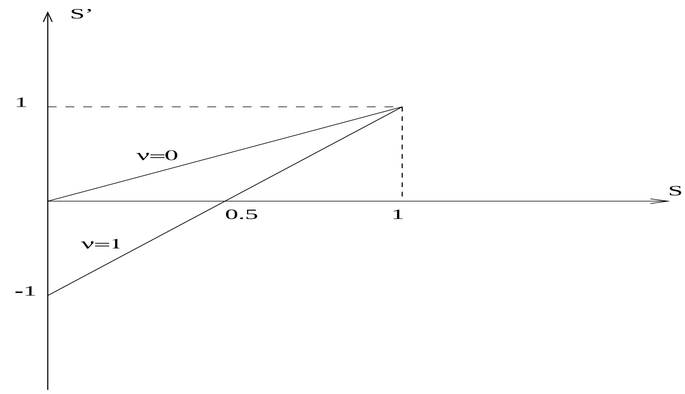


Figure 8.7: The relation between the original and modified admission policies

Therefore,

$$S'' = S \cdot (\frac{1+\nu}{\nu}) \tag{8.24}$$

Here, we have a new performance measure that should be evaluated, that is the probability of call termination, i.e.

$$C_T = \frac{\text{number of terminated calls}}{\text{total accepted calls}} \tag{8.25}$$

8.5 Statistical Analysis

8.5.1 Confidence Interval

At the end of each simulation run, we have an ensample of data (results of various criteria) for each performance measure, say $x(i)$. These ensembles shall be averaged to obtain $\overline{X}$ (the sample mean), i.e.

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x(i) \tag{8.26}$$

where n (i.e the number of samples) is determined by the required confidence interval. The sample standard deviation S is given by

$$S^2 = \sum_{i=2}^{n} \frac{(x(i) - \overline{X}\,)^2}{n-1} \tag{8.27}$$

Table 8.2: Simulation set up

Simulation equivalent period	60 min
Measurement window size	150, 200, 250, 300,350 packets
# of simulation runs	100
Time out threshold	3 windows
Reserved servers for interactive users	80%
Congestion threshold	90%

Assume that $n \geq 30$, then by applying the central limit theory, we can assert that these samples has approximately a normal distribution [101]. Now, the question is when should we stop generating new values (having more runs) assuming that we want to be at least 95% certain that our estimated parameter will not differ from the true value by more than d? From the above assumption that these generated values have a normal distribution, then it is easy to show that we should continue to generate new values until we have generated n values for which

$$\frac{1.95S}{\sqrt{n}} < d \tag{8.28}$$

When we apply the rule in Eq. (8.28) in our simulation, we observe the following. The necessary number of runs required to satisfy the condition in Eq. (8.28) widely varies depending on the traffic load as well as the estimated parameter and performance criteria. For low traffic load, we need a huge number (rage of 10^4) of simulation runs to be executed, while we need much less in the case of high traffic load. However, when we closely examine the samples collected from each simulation run, we find that the sample mean converges for most of the performance criteria especially under heavy traffic load as shown in Figs. 8.8 & 8.9. Hence, we decide to limit our simulation runs to 100 runs. Table 8.2 summarizes the simulation set up parameters. We note that during slot-II, interactive users can at most use 80% of the downlink servers. In other words, in the case of contention, we keep at least 20% of down link servers available for excess traffic packets. For the sake of practicality

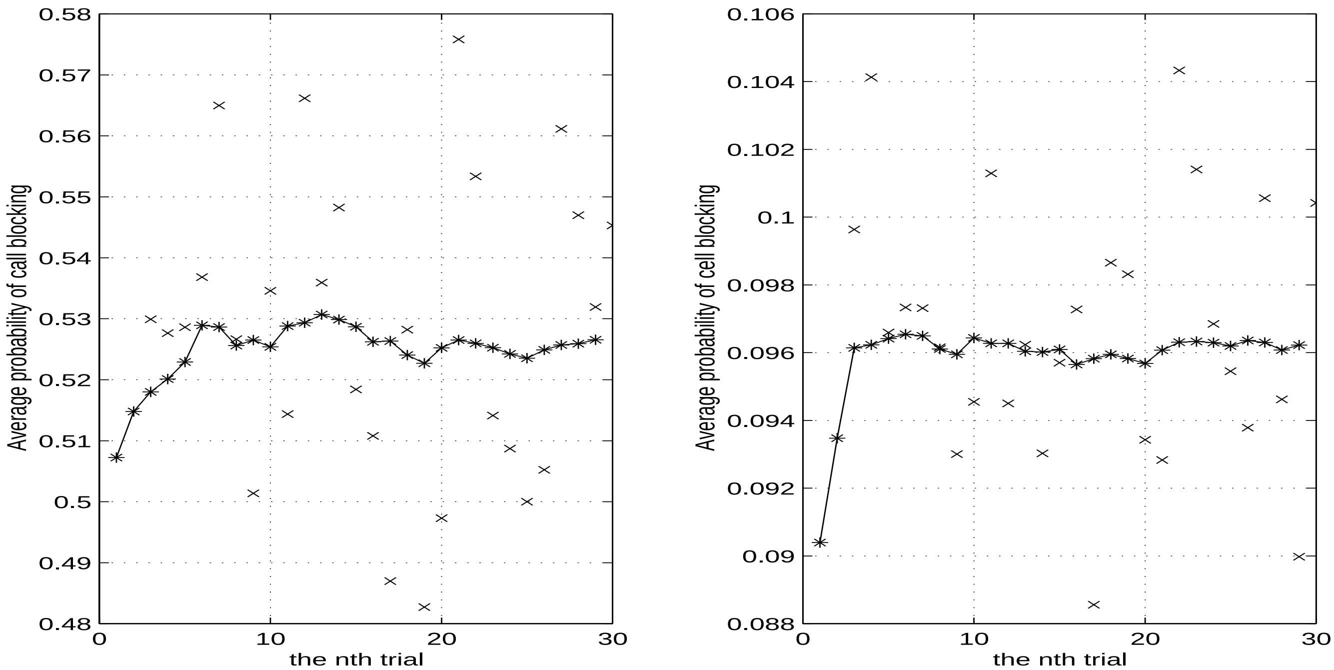


Figure 8.8: Simulation ensembles for high traffic load, SD(call)= 0.0243, SD(packet)= 0.0038

of the proposed policies, we relax the congestion threshold to be 90% instead of 100% as originally proposed. This means that the adaptation process will only be initiated if $S(.) \leq 0.9$.

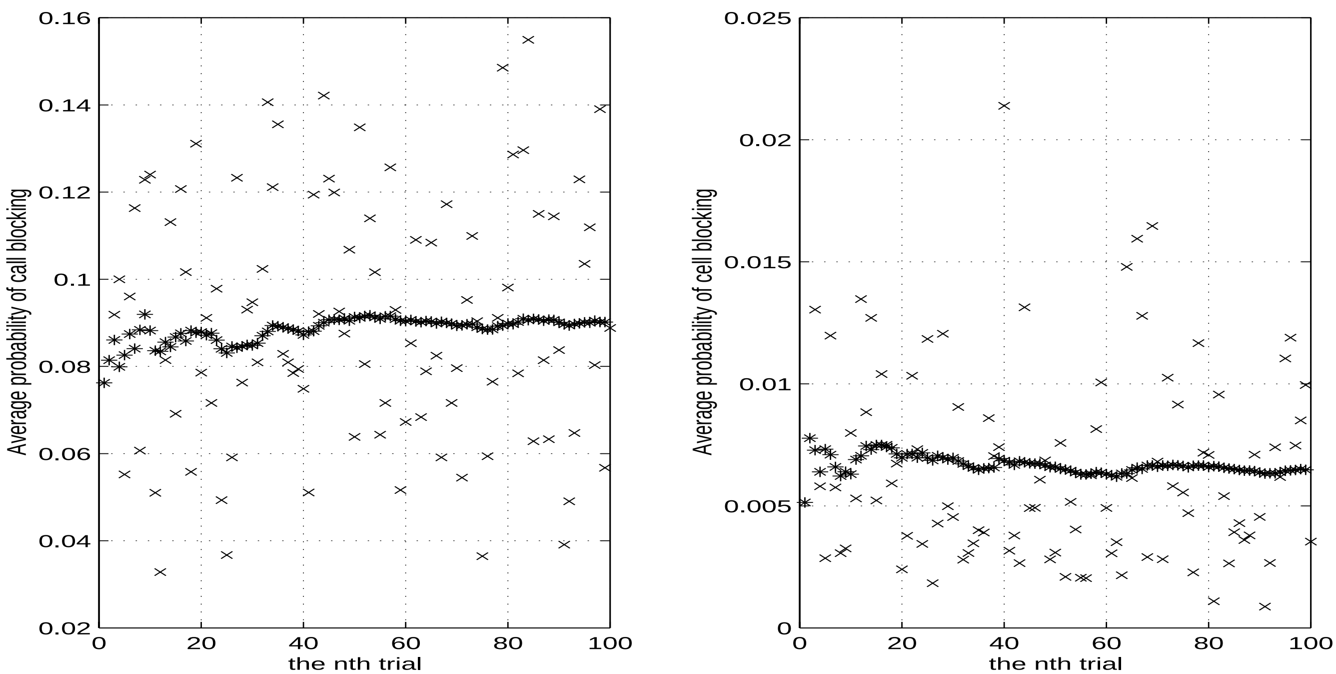


Figure 8.9: Simulation ensembles for high traffic load, SD(call)= 0.0282, SD(packet)= 0.0039

8.6 Results and Discussion

Having configured the simulation setup for our problem, we have carried out many simulation experiments to explore the system performance of the considered network under wide range of traffic load and system parameters. It is worth to emphasis that several modifications have been added to the proposed admission/Congestion policy explained in chapter 7. These modifications are adopted such that we have a more practical system.

First, a new call will be admitted if two conditions are satisfied; that are the buffer is in a good condition as well as there are enough resources to accommodate this user at its peak rate at the end of previous window. Second, during the following measurement window, the probability of congestion should be measured and consequently the current traffic will be adapted by the adjusting the traffic load that has been admitted during the very last measurement window. As we have mentioned in section 8.4.2, we are only going to control the traffic portion that was recently admitted in the last window. Third, several scenarios of the proposed admission/congestion policy are simulated. Fourth, the call duration has been explicitly simulated to be geometrically distributed. In contrast, in the analytical analysis, the call duration was not taken explicitly into consideration. However, we have studied the buffer activity under all packet generation possibilities as was explained in chapter 7.

This adaptation for those newly admitted users can be done through one of the following ways: *termination*, *bite rate variation* or both. The former way means that some or all new calls should be terminated to maintain certain QoS requirements. On the other hand, a new user may be requested to lower his transmission bit rate such that the system performance under the new load could be acceptable. It is obvious that we could have many scenarios and the matter of choosing which should be applied depends either on the kind of service offered by the network or on the kind of service paid for by the user. Here, however, we shall apply the *termination* method in controlling the traffic load.

Further, both policies mentioned in section 8.4.2 will be compared when it is applicable to the *Basic Admission Policy*. This policy is a primitive one and it just gives an indication about how the traffic burstiness at the base station have been smoothened after a certain period of time (i.e. window). We shall consider the system performance under this policy as a reference for the other two policies when it is applicable.

The first group of simulation results shall explore the network performance under different admission/congestion policies for fixed base station buffer size (i.e. $B = 80$) and fixed measurement window size (i.e. $W = 150$).

Figures 8.10-8.12 show the average call blocking probability for stream traffic users under different interactive traffic loads for the three admission/congestion policies. The higher the interactive traffic, the higher is the stream call blocking probability. This is a natural result of the proposed hybrid MC-TDMA/CDMA, because many resources (i.e. receivers) are reserved for high bit rate interactive users even though their activities are low. Hence, many stream call requests will be blocked because the excess traffic could not be handled together with interactive traffic on slot-II.

Therefore, to alleviate this problem, resources should not be reserved for high bursty users depending on worst case (i.e. peak transmission rate), but on other schemes such as average bit rate where smart queuing is provided at the user's site. This conclusion is also supported by the results shown in chapter 7, where the CDMA traffic is modeled as M/M/1 system assuming an infinite buffer capacity at the user site.

More, the other factor that enhances the call blocking probability is the call termination policy. As we can observe from the system performance under policy-I and policy-II, the call termination is high especially under low interactive traffic and high stream traffic. For instance, about 15% of accepted call have been terminated under policy-II (under $\rho_{Int} = 6.13 Erlang,\ \ \rho_s = 55.21 Erlang$), while more than 60% have been terminated under policy-II. On the other hand, policy-III causes no single call to be terminated.

Nevertheless, it is worth to emphasis that during the course of simulation, we assume that all currently admitted users are requesting the same QoS (i.e. buffer not congested for q consecutive frames). In practice, however, it is not necessary that all users ask for the same QoS. Consequently, the call termination probability will be less. Alternatively, the rejection action can be relaxed by allowing the system performance to degrade for a longer period of time (e.g. 3 windows) hopping the traffic shall statistically be multiplexed. On the other hand, the network operator may choose to adapt the transmission rate (traffic shaping) of some users according to certain flow control policy instead of rejecting the calls completely (e.g. [85]). Here, of course, it is assumed that the end user can support the large volume of burst packets.

Figures 8.13-8.15 illustrate some advantages that the stream users will gain as a result of the high price paid in call blocking probability. First, the packet drop under policy-I is very low compared to the performance under policy-II or policy-III. Industry's standards (e.g. IEEE 802.16.1) request that the packet blocking should not exceed 1%. Considering this specification, we find that policy-I can guarantee this QoS for up to 40 Erlang stream traffic load under a wide range of interactive load. On the other hand, policy-II and policy-III can support up to 33 and 30 Erlang stream loads, respectively.

Second, policy-I has effectively limited the packet delay for stream users. Under the same traffic load, policy-I limits the packet delay to one third of the delay under policy-III and to one half the delay experienced by the stream packets under policy-II. Third, the end-to-end packet losses is guaranteed to be low under all proposed policies as shown in Figs. 8.16-8.18. Moreover, Policy-I offers the least level of packet losses among the three policies. This feature is very attractive for stream

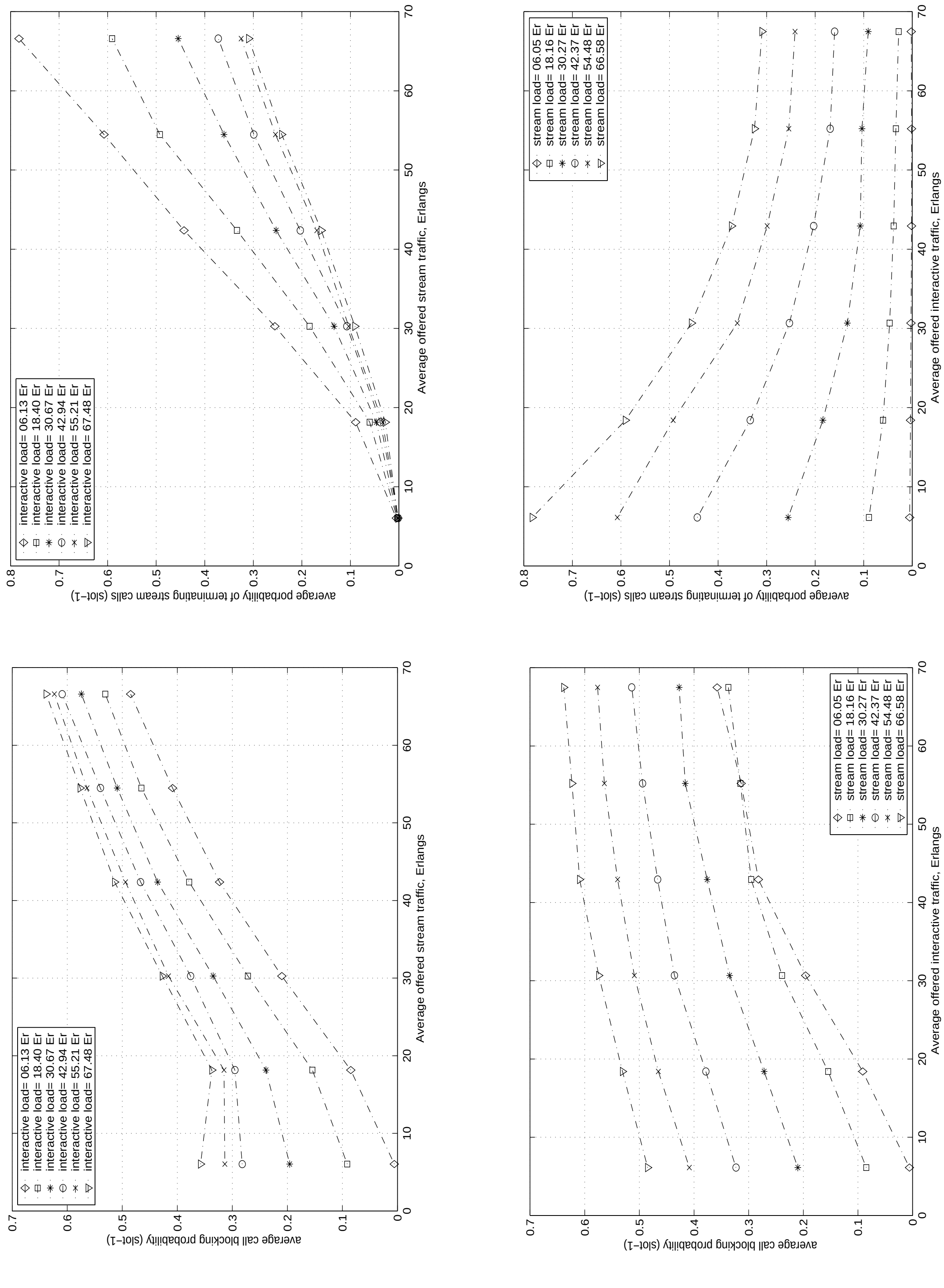


Figure 8.10: The comparison of simulation results of call blocking and call termination probabilities under policy-I, where $W = 150$ packets, $q = 90$, and buffer capacity= 80

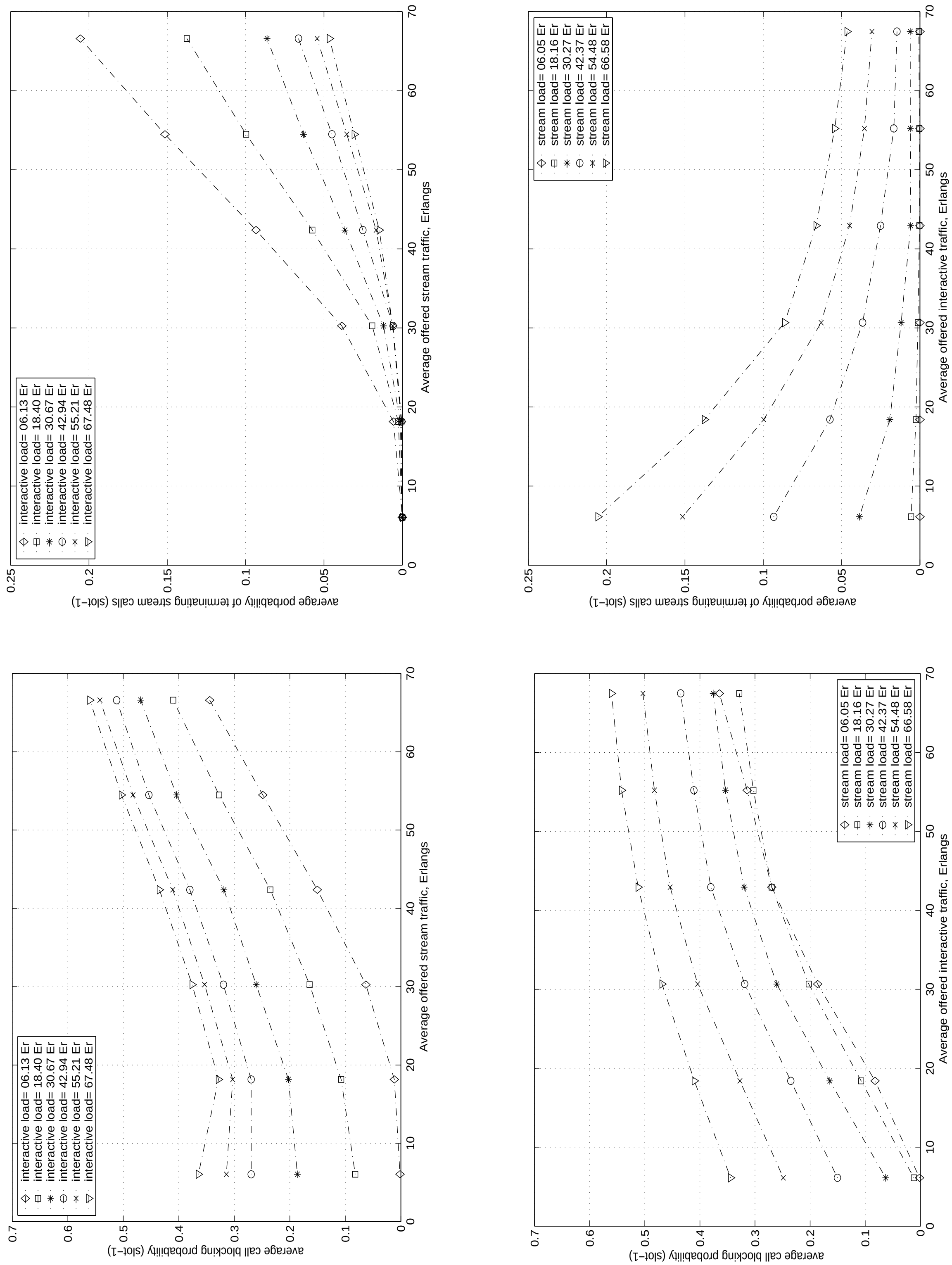


Figure 8.11: The comparison of simulation results of call blocking and call termination probabilities under policy-II, where $W = 150$ packets, $q = 90$, and buffer capacity= 80

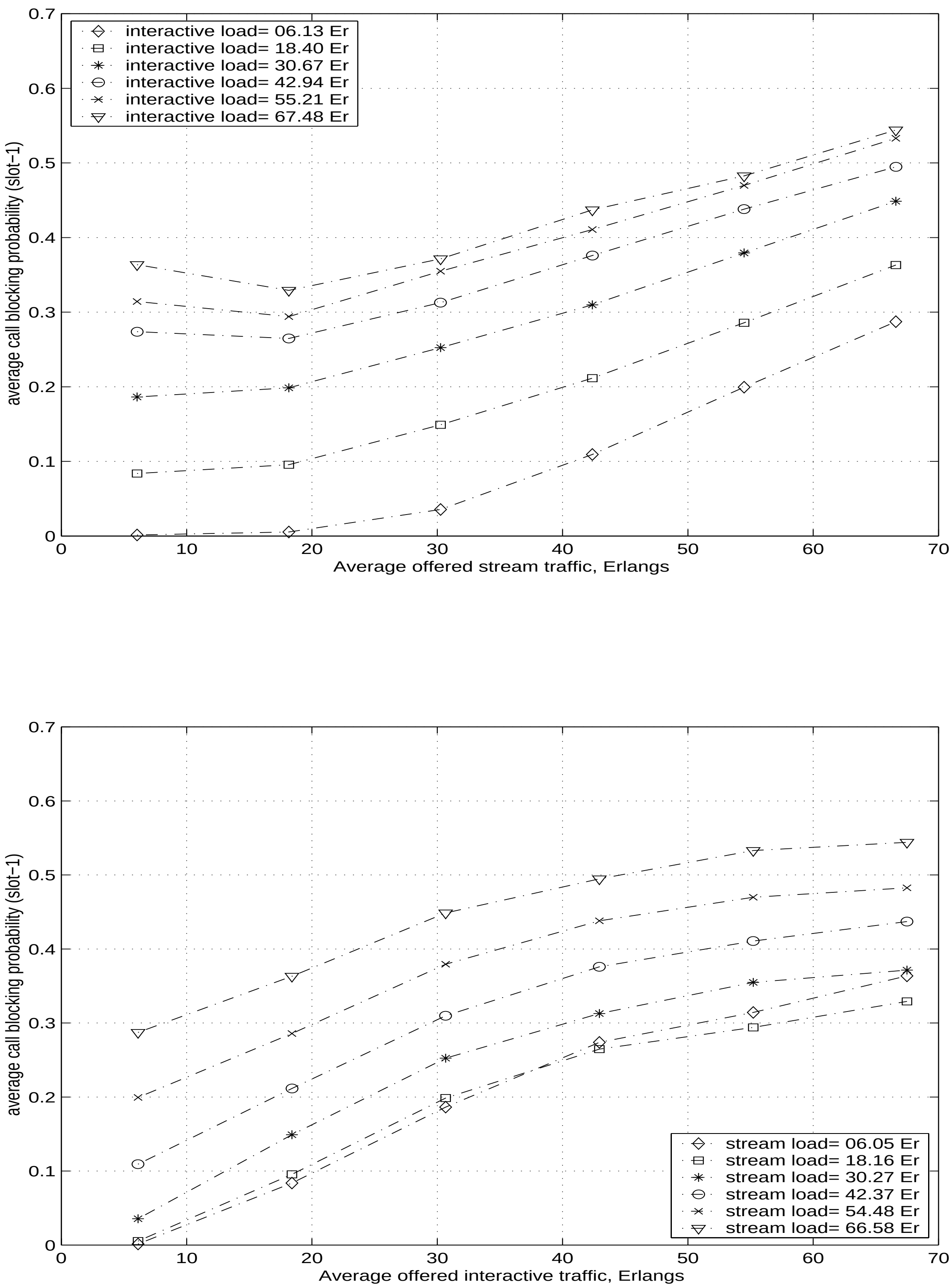


Figure 8.12: The comparison of simulation results of call blocking probability under policy-III, where $W = 150$ packets, $q = 90$, and buffer capacity= 80

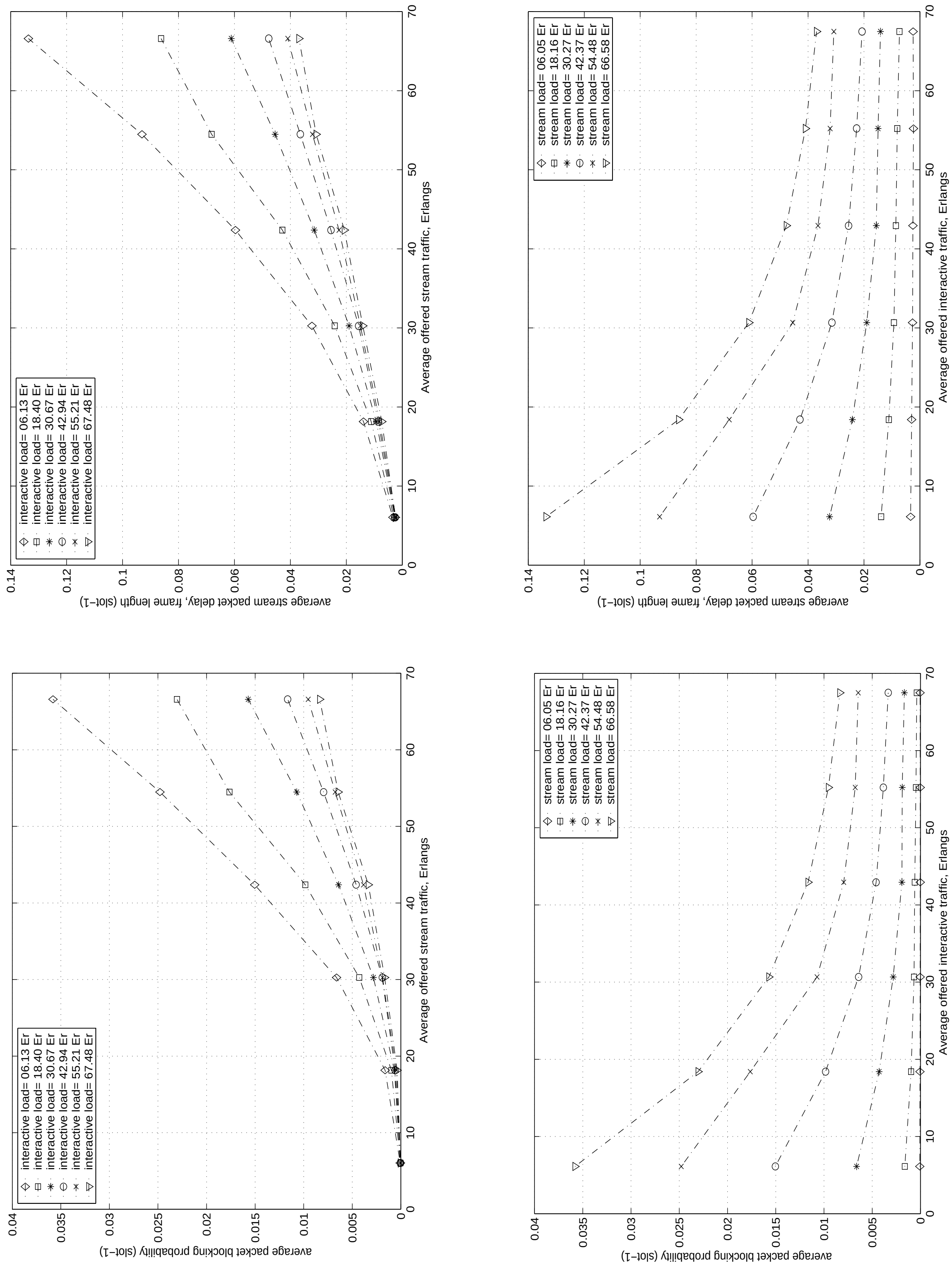


Figure 8.13: The comparison of simulation results of packet blocking probability and packet delay under policy-I, where $W = 150$ packets, $q = 90$, and buffer capacity= 80

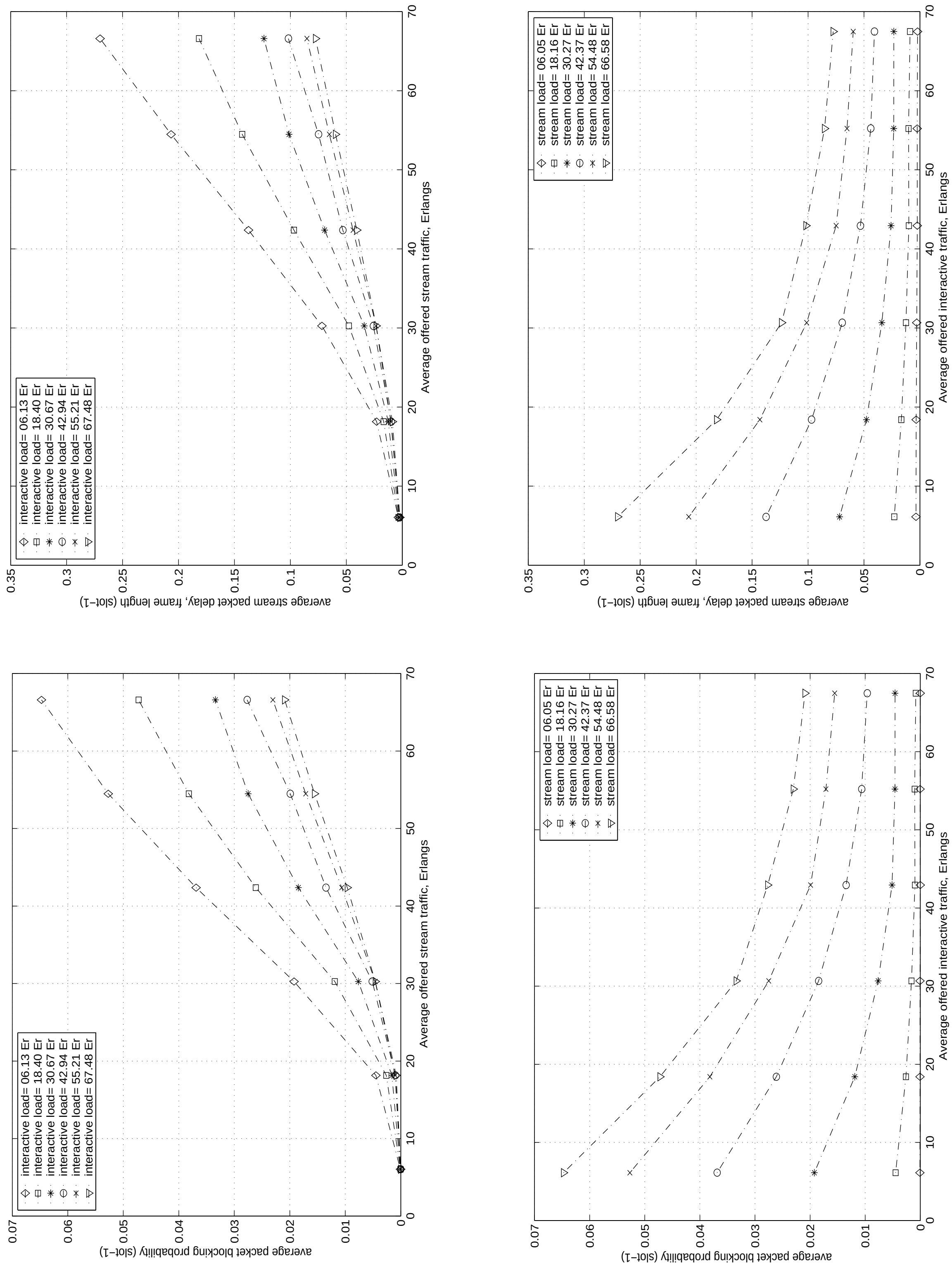


Figure 8.14: The comparison of simulation results of packet blocking probability and packet delay under policy-II, where $W = 150$ packets, $q = 90$, and buffer capacity= 80

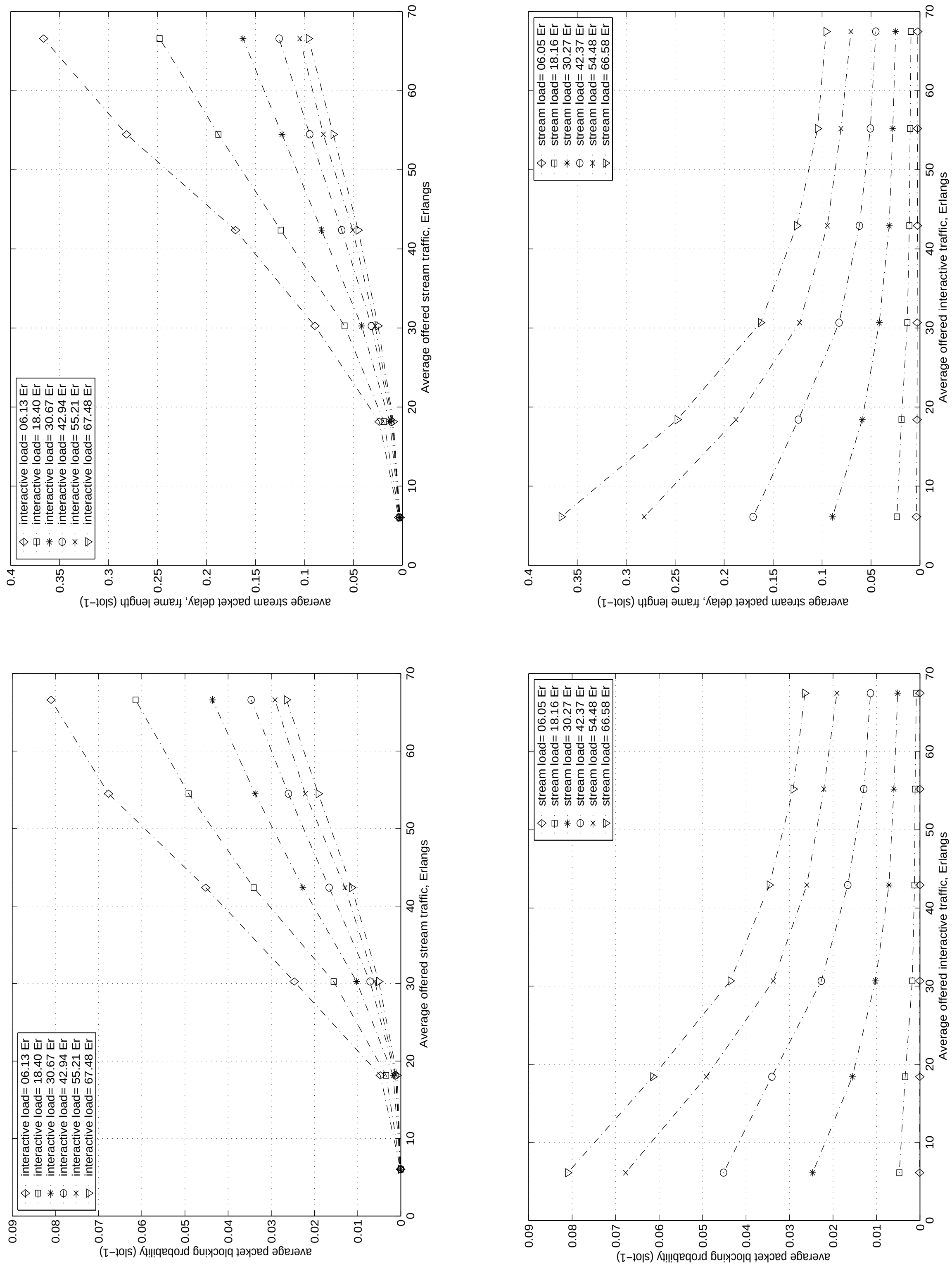


Figure 8.15: The comparison of simulation results of packet blocking probability and packet delay under policy-III, where $W = 150$ packets, $q = 90$, and buffer capacity= 80

traffic applications. Yet, policy-I shows a low system throughput compared with the two policies. This is due to high call termination under this policy which is imposed to guarantee a very stringent QoS that is the active stream should find the buffer not congested for a consecutive q packet-time units.

Figures 8.19-8.21 depict the probability of non-congested buffer under the three policies. Here, it is not fair to compare these policies to each other for this parameter of performance because each one is evaluated under different scale. However, we notice the following. Although, policy-I has been adopted to guarantee that this performance measure should not fall below 90%, it fails to maintained this QoS under heavy interactive load.

On the other hand, policy-II does guarantee the requested QoS (i.e. Prob. Of non-congested buffer $\leq 90\%$). Also, when the probability of non-congested buffer under policy-III has been evaluated in similar scale as policy-II, it shows that this QoS can be maintained to be less than 90%. This phenomenon can be attributed to the rejection policy where we just reject the very recently admitted calls, while in fact, a portion of old admitted calls should be adapted also.

Considering the performance of interactive users along the excess stream traffic using slot-II, the simulation results show no buffer congestion, and very small packet delay. However, the two performance parameters that should be examined are the call blocking probability of interactive users requests and the end-to-end interactive packets losses, because these measures have a great impact of the quality of these services. Figures 8.22- 8.24 illustrate the call blocking and packet loss probabilities under the three proposed policies. First, we observe that the stream traffic load plays a little role regarding these parameters especially under policy-I. The reason for that is the network design where the interactive callers have granted higher priority than the stream callers. The call blocking is too high, since the admission policy reserves resources for active interactive users depending on their peak rates. Second, the packet losses are low and satisfactory.

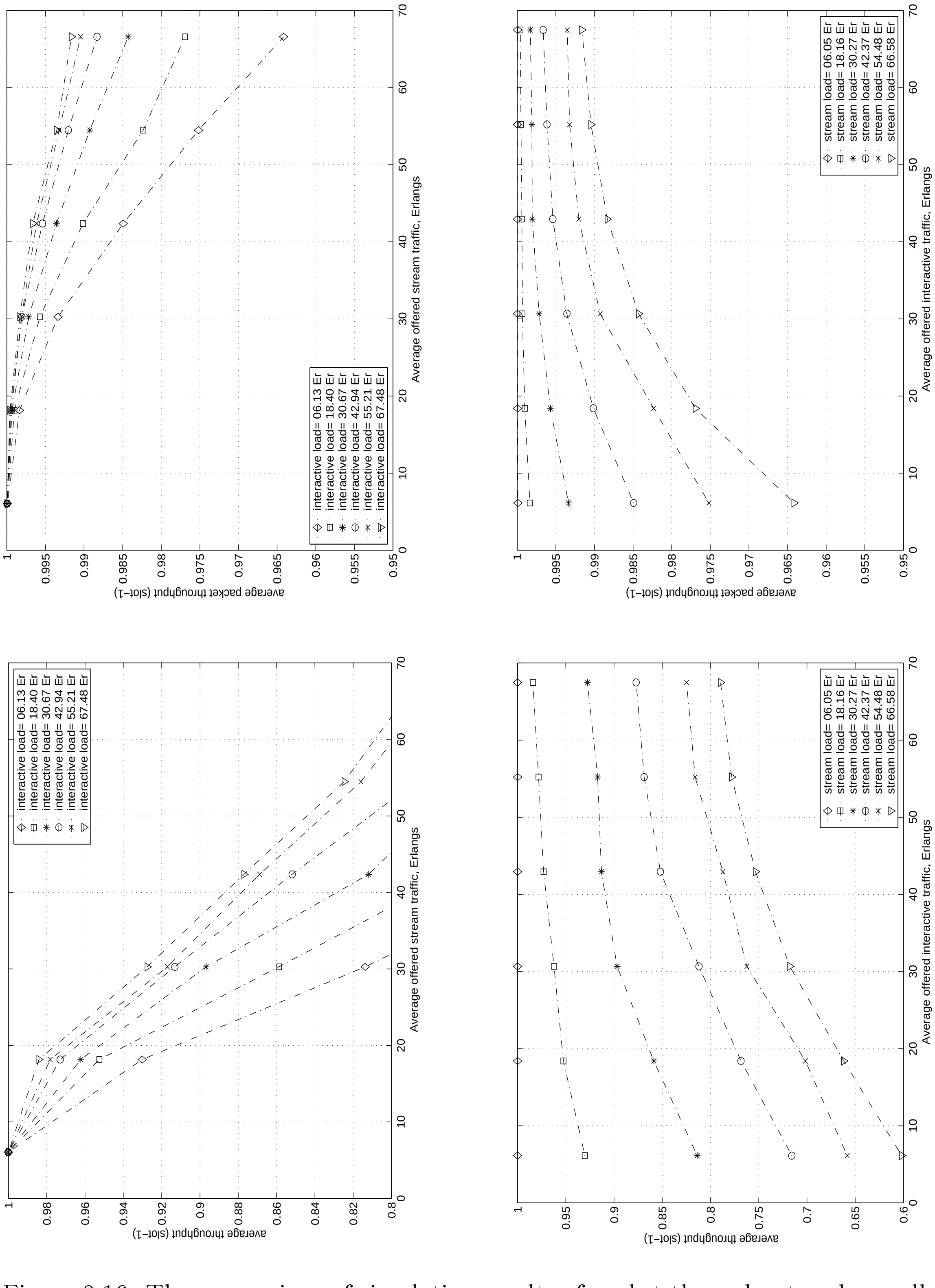


Figure 8.16: The comparison of simulation results of packet throughput and overall packet loss under policy-I, where $W = 150$ packets, $q = 90$, and buffer capacity= 80

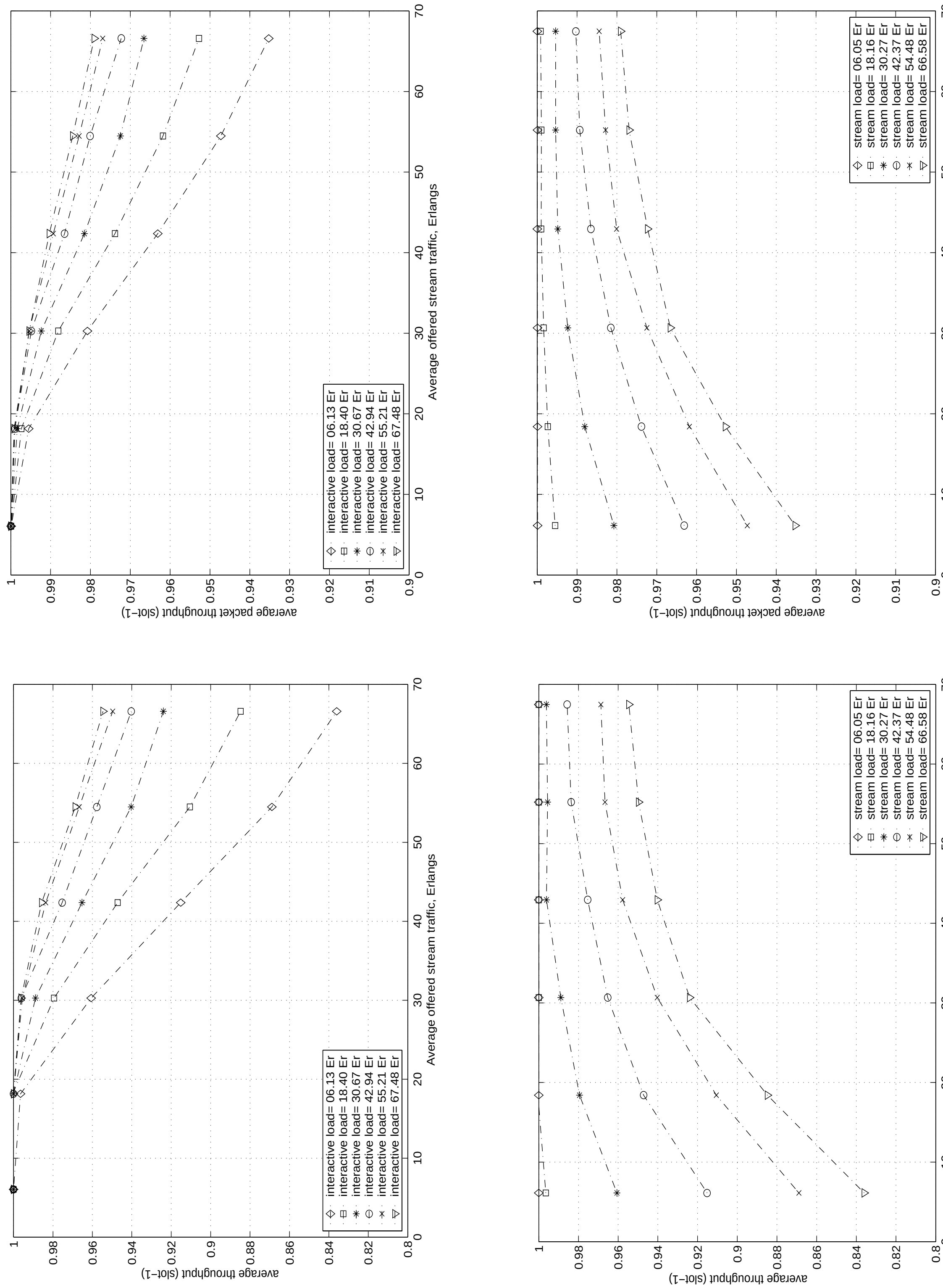


Figure 8.17: The comparison of simulation results of packet throughput and overall packet loss under policy-II, where $W = 150$ packets, $q = 90$, and buffer capacity= 80

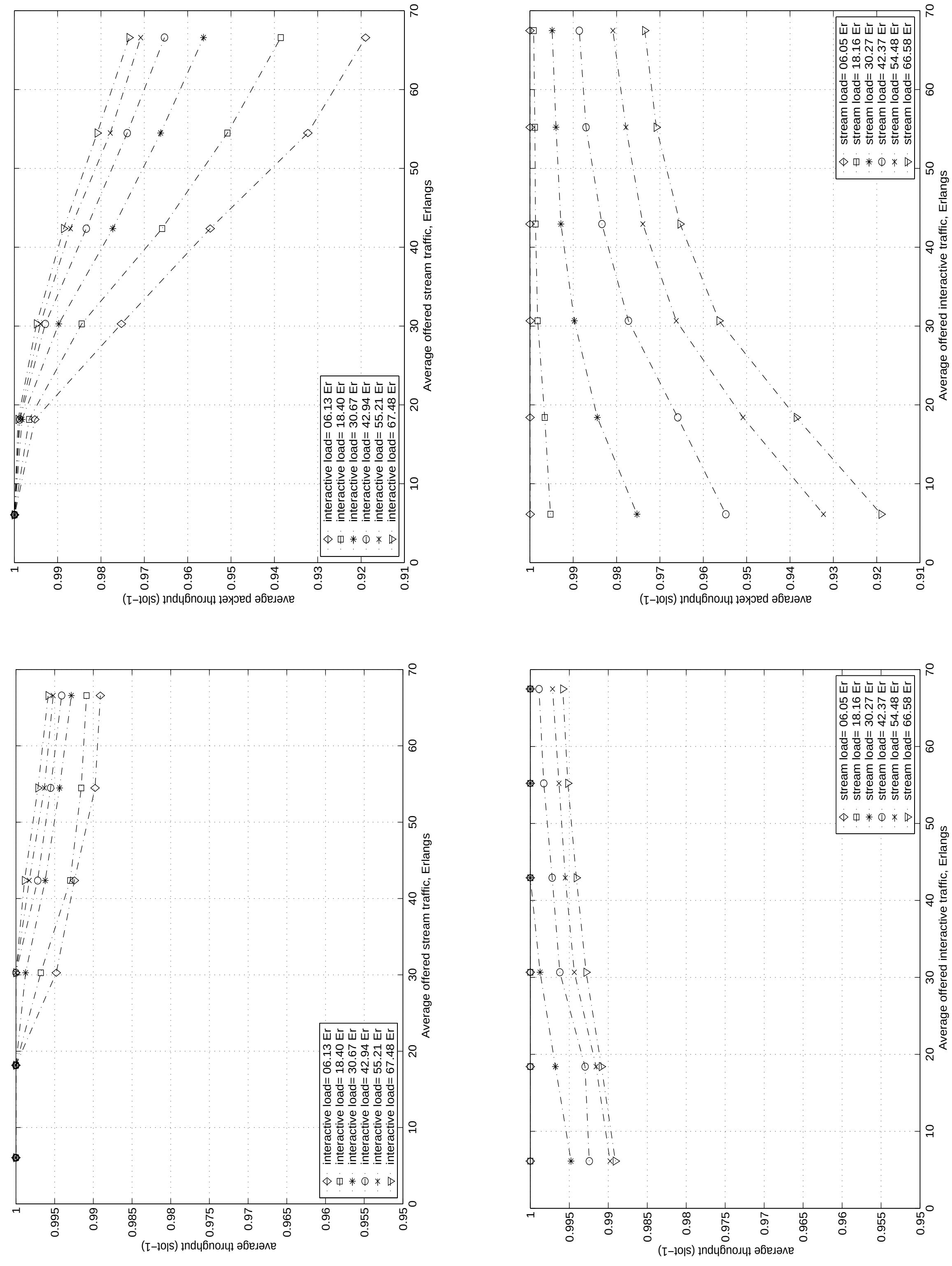


Figure 8.18: The comparison of simulation results of packet throughput and overall packet loss under policy-III, where $W = 150$ packets, $q = 90$, and buffer capacity= 80

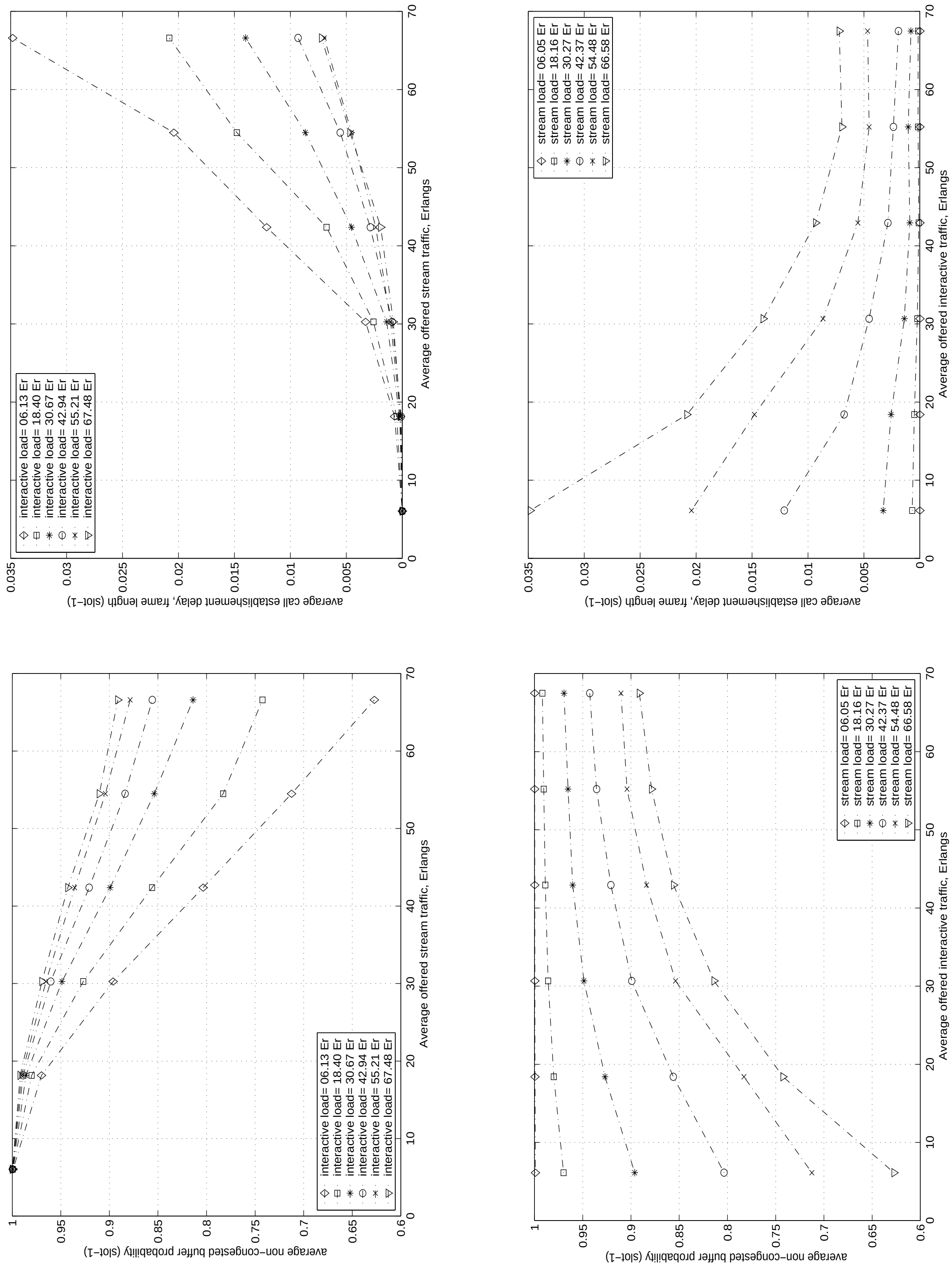


Figure 8.19: The comparison of simulation results of the probability of non-congested buffer and average call establishment delay under policy-I, where $W = 150$ packets, $q = 90$, and buffer capacity= 80

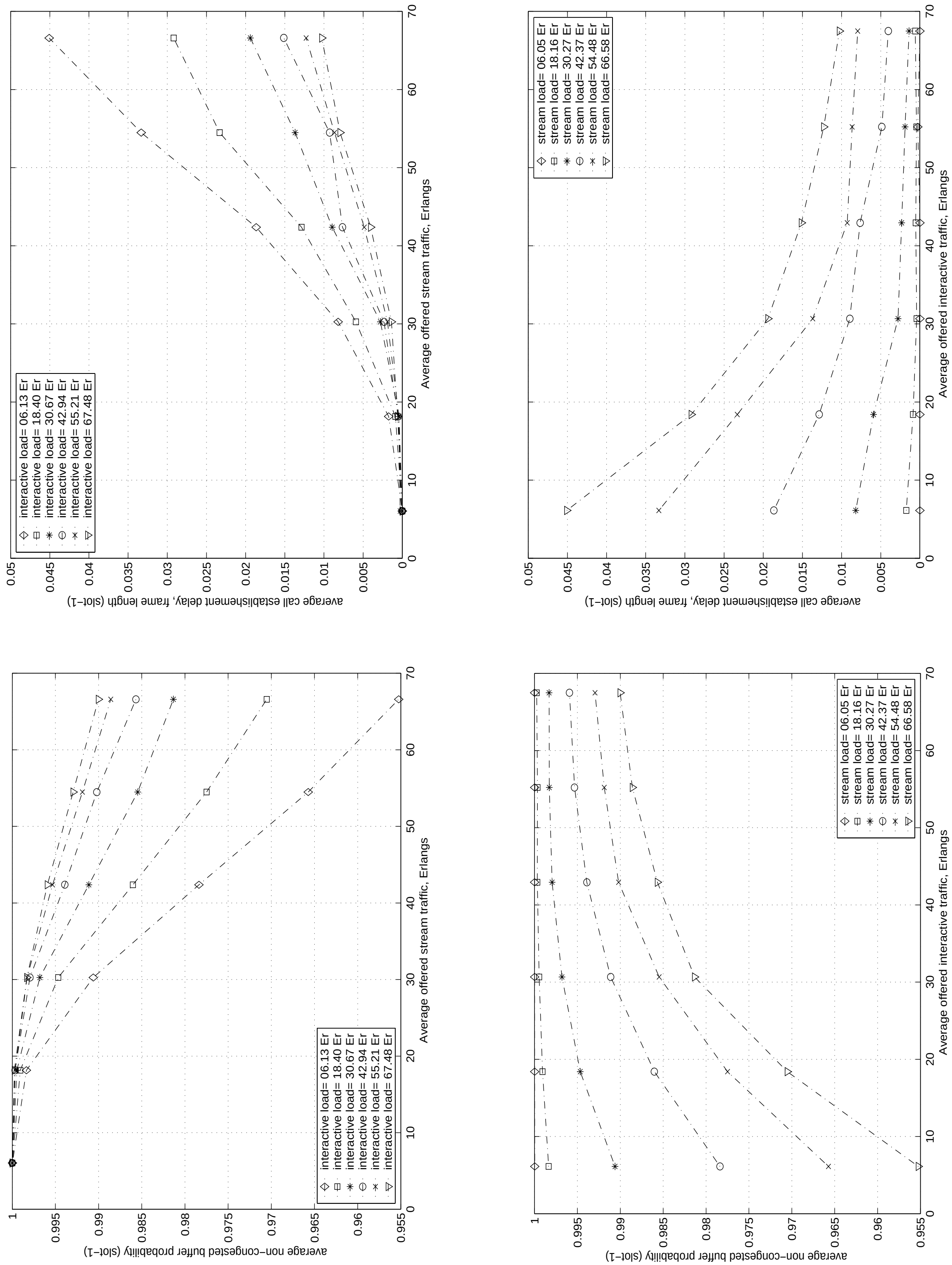


Figure 8.20: The comparison of simulation results of the probability of non-congested buffer and average call establishment delay under policy-II, where $W = 150$ packets, $q = 90$, and buffer capacity$= 80$

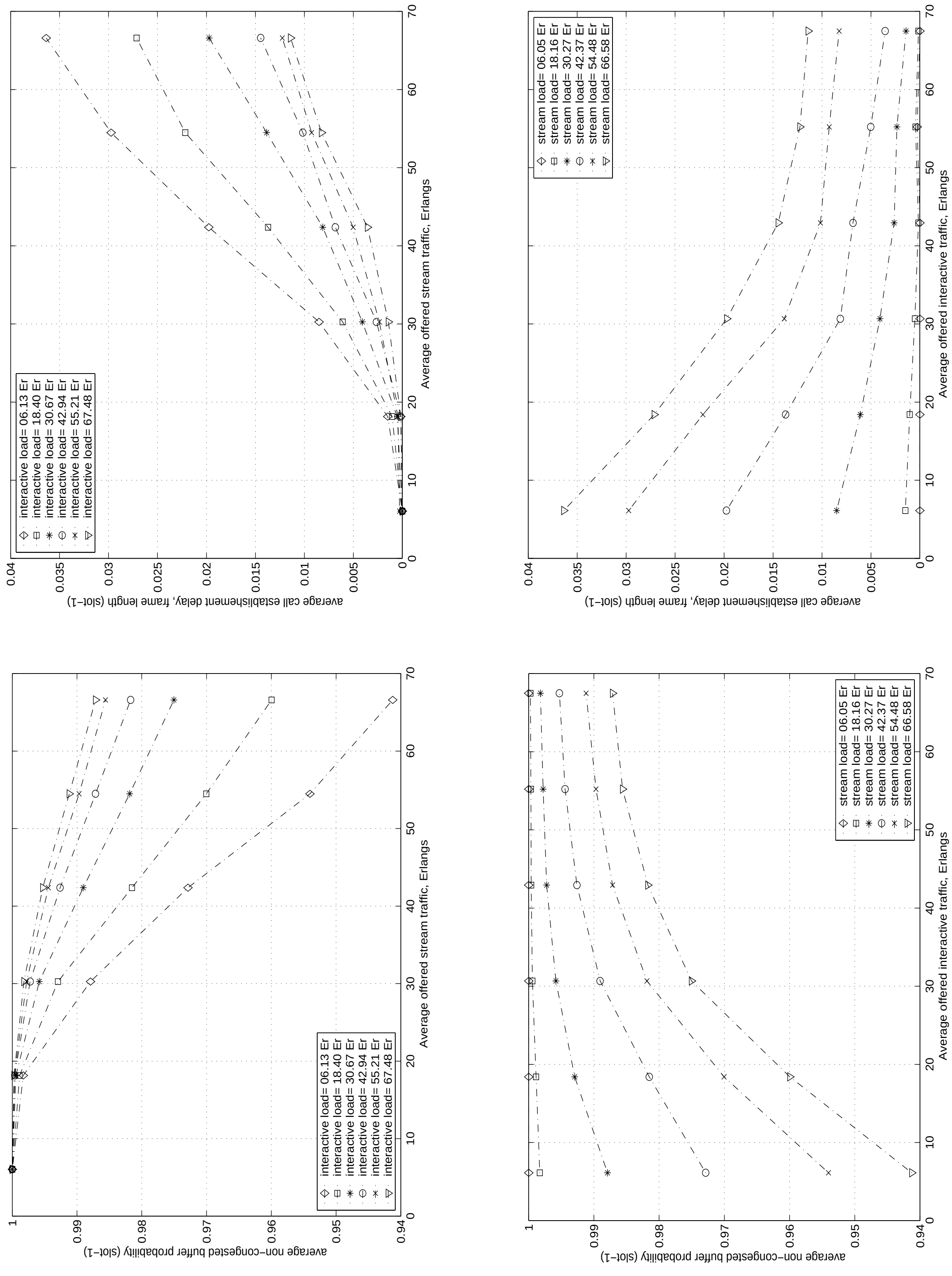


Figure 8.21: The comparison of simulation results of the probability of non-congested buffer and average call establishment delay under policy-III, where $W = 150$ packets, $q = 90$, and buffer capacity= 80

Figure 8.25 shows the packet error rate under the three policies for slot-I traffic. Policy-I outperforms the other two, because policy-I admits lower traffic to the network. On the other hand, we find the uplink packet error is very low because the decorrelator receivers are used for signal detection which becomes more practical by applying our proposed flow traffic control. Considering the same performance measures for the traffic population using slot-II, we observe these measures have been maintained for both links as shown in Fig. 8.26. More, it is important to note that maintaining the packet error within the acceptable range is a byproduct of our proposed admission policy. It proves that including this parameter in the queueing problem does result in maintaining this QoS even though, our admission/congestion policy does not target it directly.

Having discussed the network performance under different admission/congestion policies where all depend in their decision on the status of the buffer at the base station over a fixed period of time (i.e. measurement window). Now, we want to study how does the variations in the measurement window size W affect the network performance?

First, we consider the call blocking probability. We notice that the larger the measurement window, the higher is the call blocking probability. This can be attributed to the fact that when we enlarge the window size, the decision making might be late and in the mean time more new call requests have initiated and then more calls will be terminated as it is shown clearly in Figs. 8.27-8.29. Hence, these terminated calls actually contributes to the total call blocking as defined above.

Also, these figures show the enhancement of packet blocking probability for larger measurement window size. Again, this phenomenon can be explained in the light of the above reasoning that when the decision making is late many received packets via uplink might find the buffer is congested as shown in Figs. 8.30-8.32. Moreover, having congested buffer leads to higher probability of unsuccessful reception on the downlink because of the huge number of simultaneously transmitted

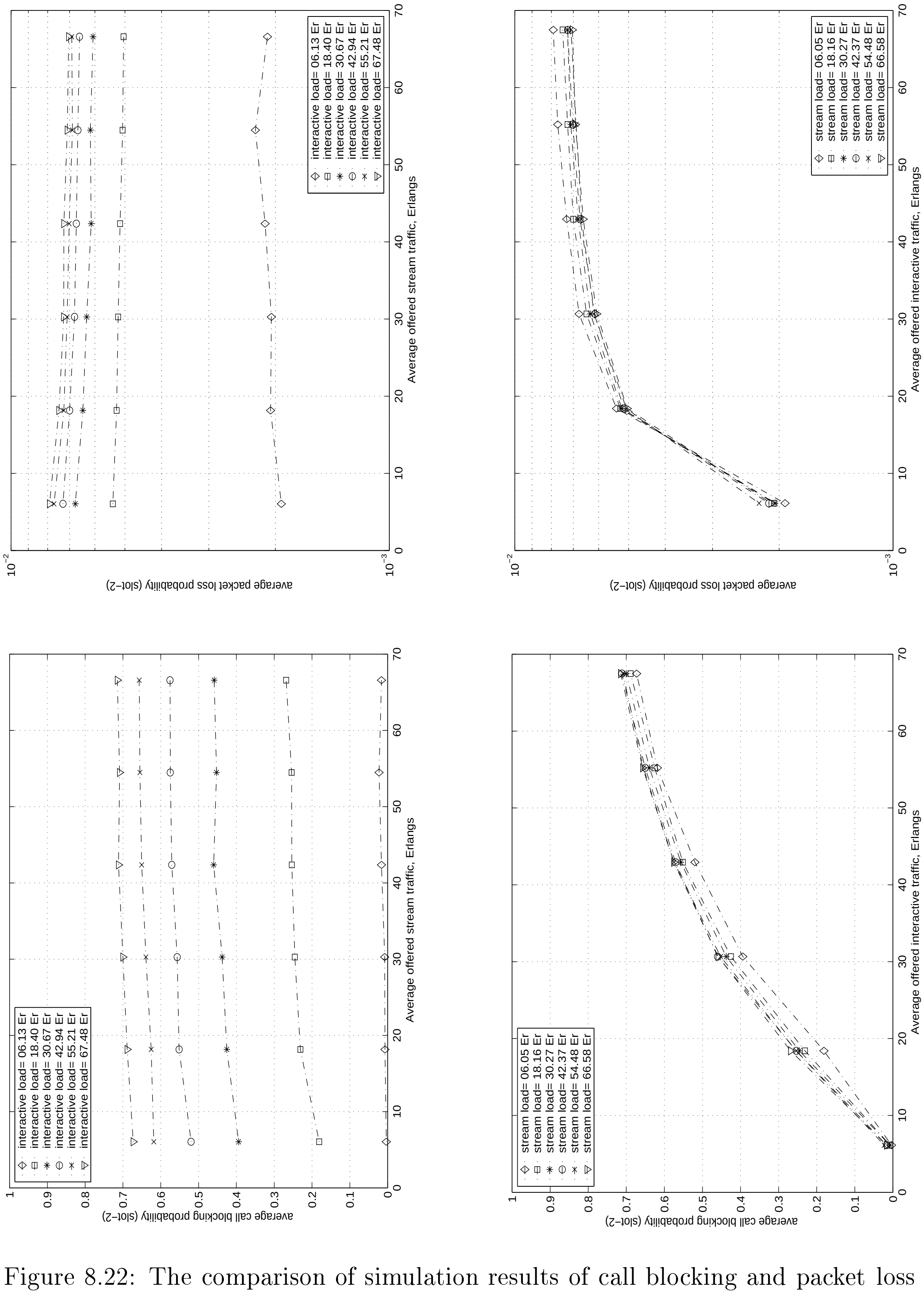


Figure 8.22: The comparison of simulation results of call blocking and packet loss probabilities under policy-I, where $W = 150$ packets, $q = 90$, and buffer capacity= 80

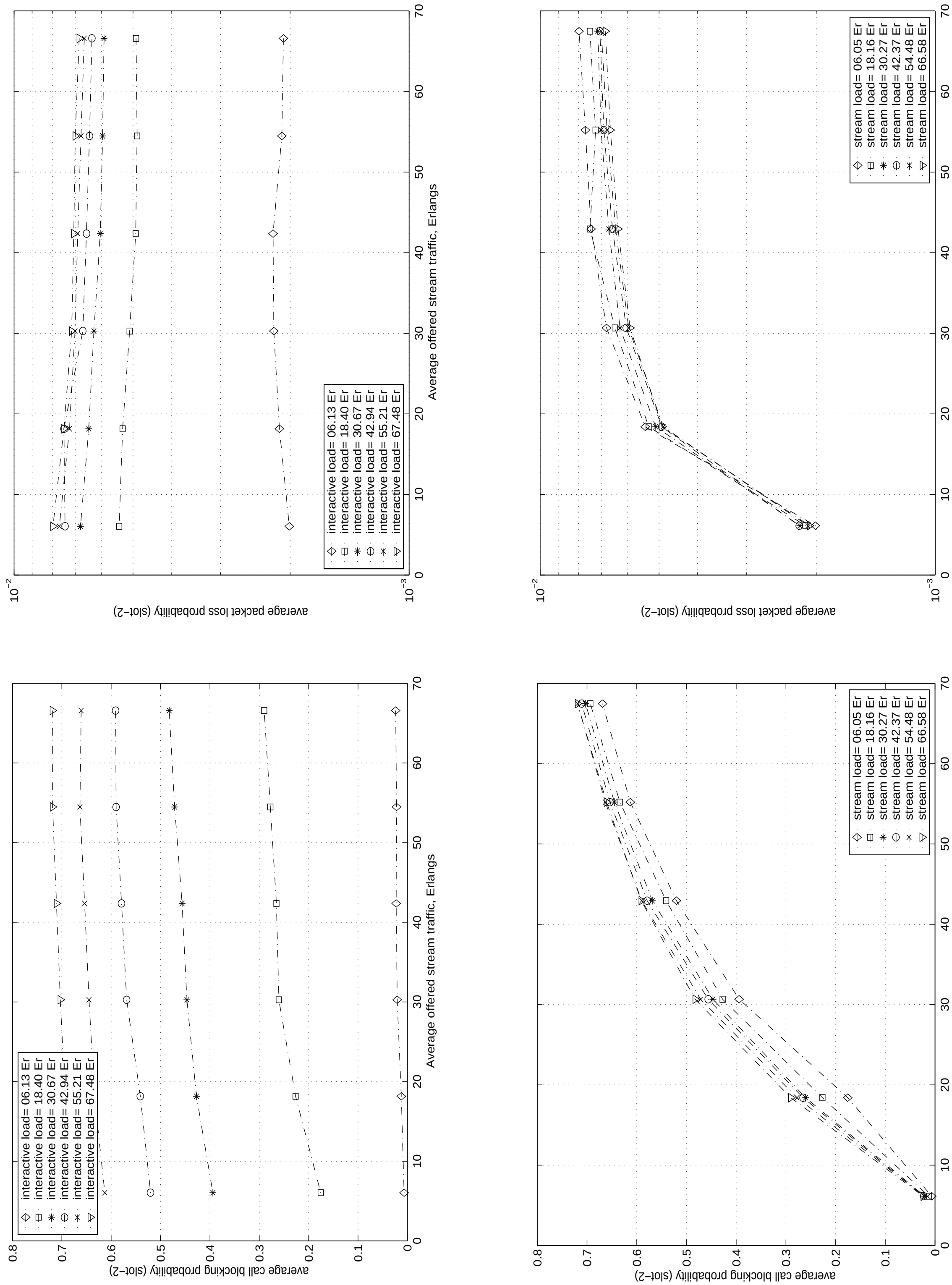


Figure 8.23: The comparison of simulation results of call blocking and packet loss probabilities under policy-II, where $W = 150$ packets, $q = 90$, and buffer capacity= 80

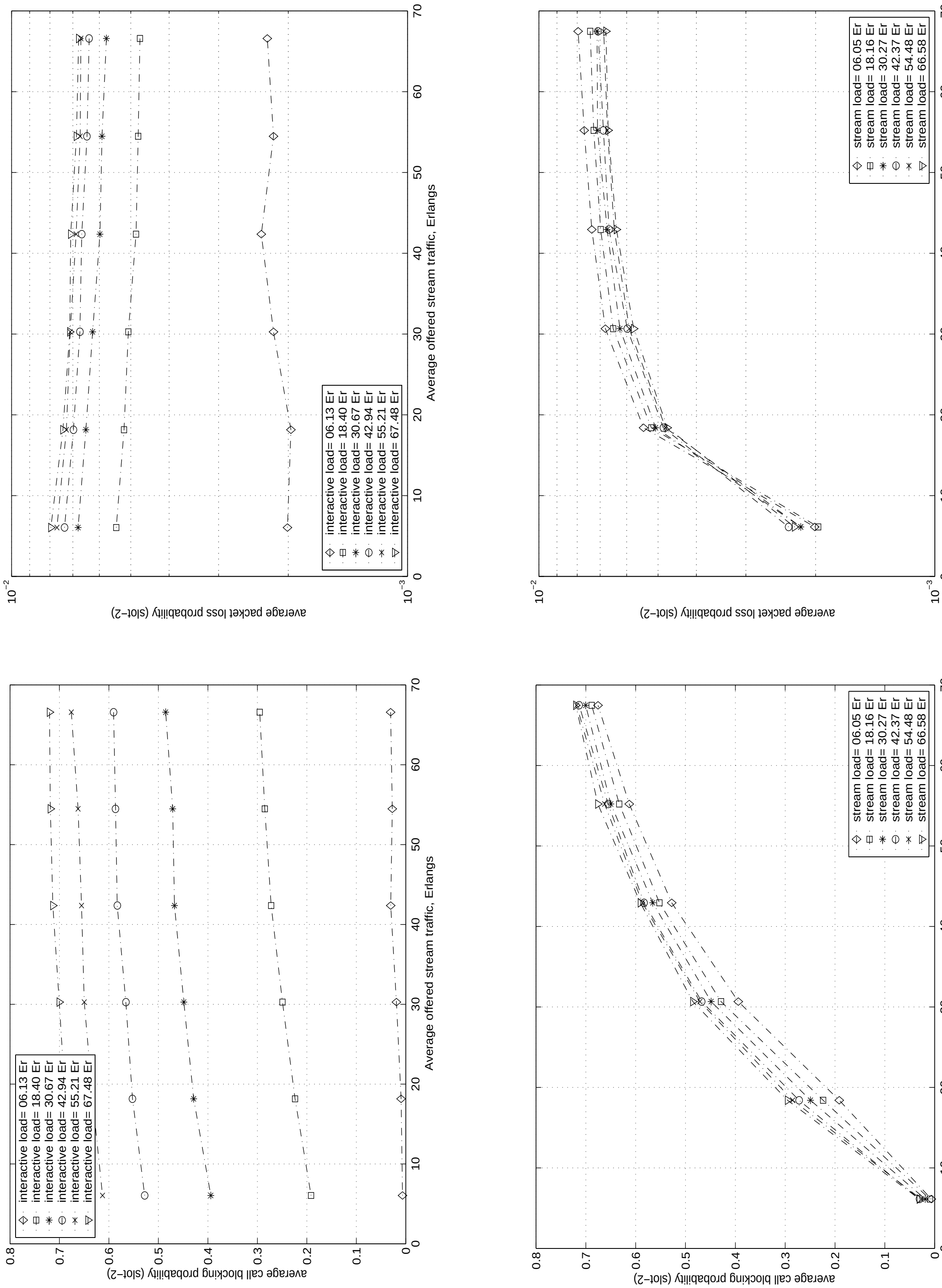


Figure 8.24: The comparison of simulation results of call blocking and packet loss probabilities under policy-III, where $W = 150$ packets, $q = 90$, and buffer capacity= 80

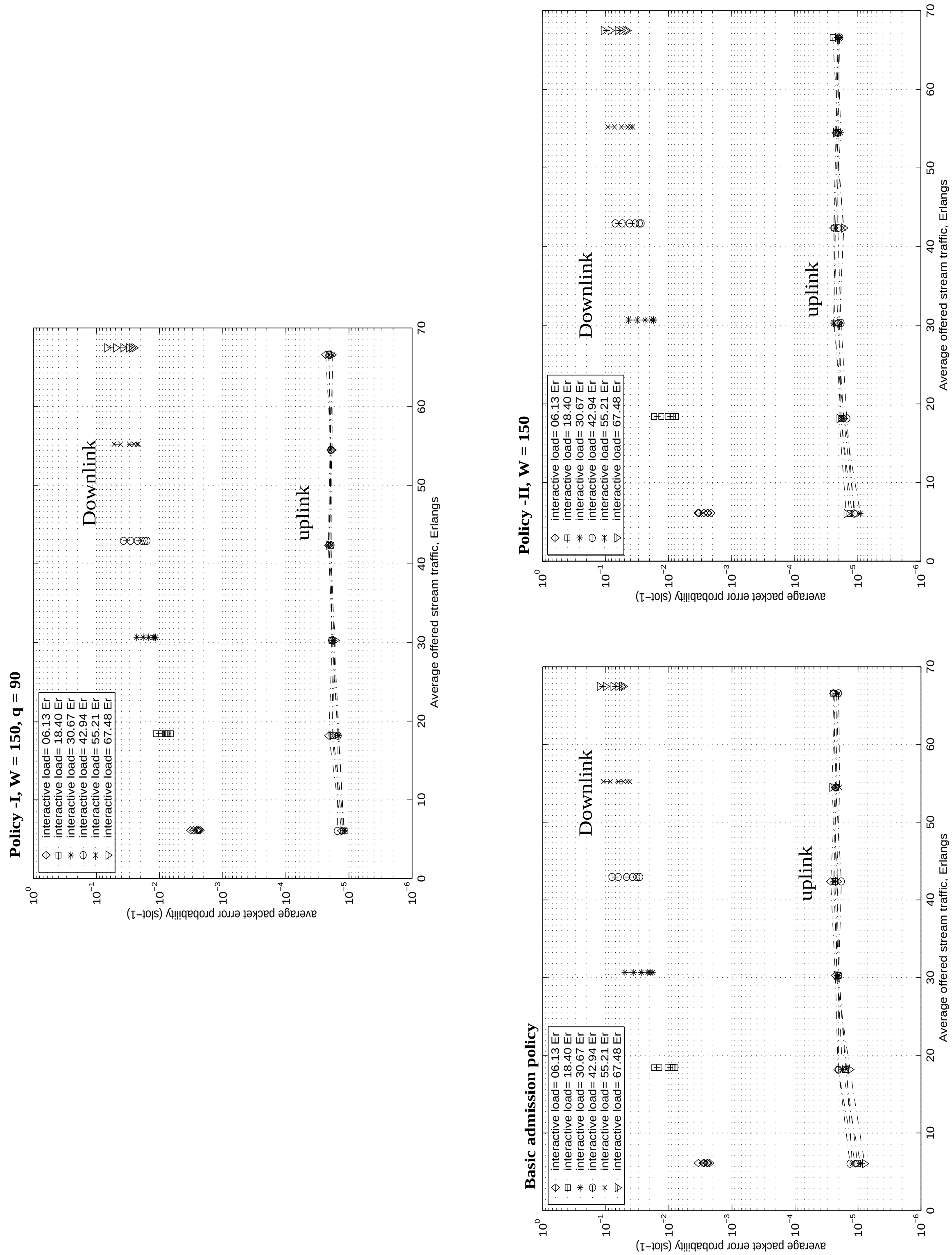


Figure 8.25: The comparison of simulation results of packet error rate under all proposed policies for traffic on slot-1, where $W = 150$ packets, $q = 90$, and buffer capacity= 80

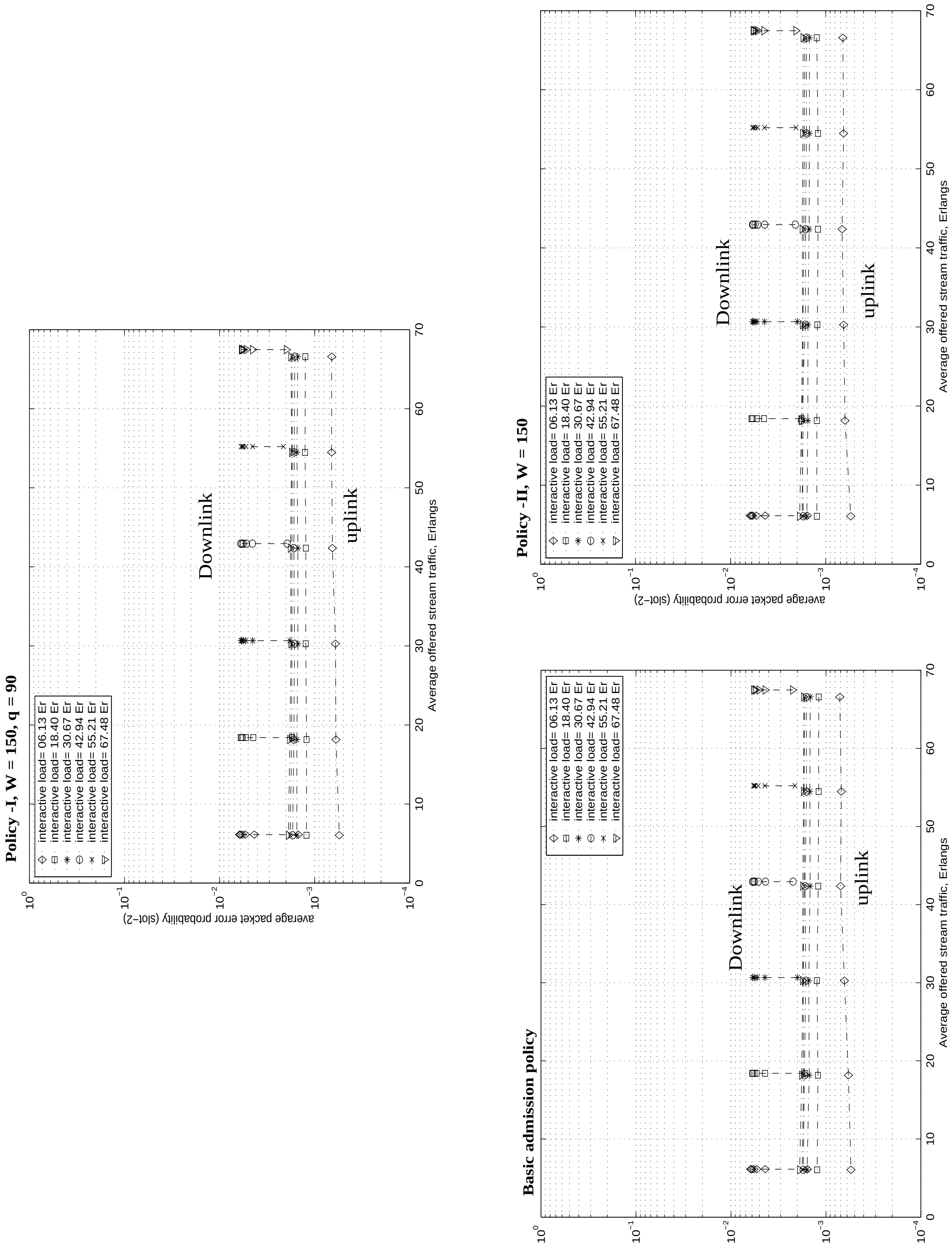


Figure 8.26: The comparison of simulation results of packet error rate under all proposed policies for traffic on slot-2, where $W = 150$ packets, $q = 90$, and buffer capacity= 80

packets. Consequently more packet should be retransmitted and eventually longer packet delay as shown clearly in Fig. 8.30-8.32. More, these figures show that the window size has little effect on the average throughput of the stream packets transmitted in Slot-I. Further, we can notice that the call establishment delay is bigger when we have larger measurement window size especially high traffic load. This phenomenon is related to the enhancement in call blocking probability for larger window size where the admission decision is postponed longer when measurement window size is larger.

Now we turn to examine the effect of measurement window size on the interactive traffic performance in Slot-II. Figures 8.27-8.29 shows lower interactive call blocking probability for larger windows which is due to the higher stream call blocking when the measurement window is large.

Considering the same performance measures under policy-2, we find that the system performance is less sensitive to the measurement window size compared to the system under policy-1. Moreover, the interesting result is that the larger the window size, the better is the packet throughput. Also, we observe, as shown in Fig. 8.31, that the probability of call terminated is lower for larger window size. These two results are due to the fact that when we have a larger measurement window we allow the traffic burstiness to be smoothened and the admission decision at the end of the window should more reliable. However, the call establishment delay increases as the measurement window size does so.

The other crucial design parameter is the base station buffer size. Figures 8.35-8.37 show little sensitivity of call blocking probability to the buffer size for the network operating under policy-1 and policy-III. However, the larger the buffer size the greater is the call blocking under policy-II. The enhancement under policy-II is about 40% if $B = 100$ compared to the case when $B = 60$. This may be attribute to high percentage of terminated calls as shown in the above figures. Of course, this is a disadvantage of offering large buffer where more packets can reside in and then

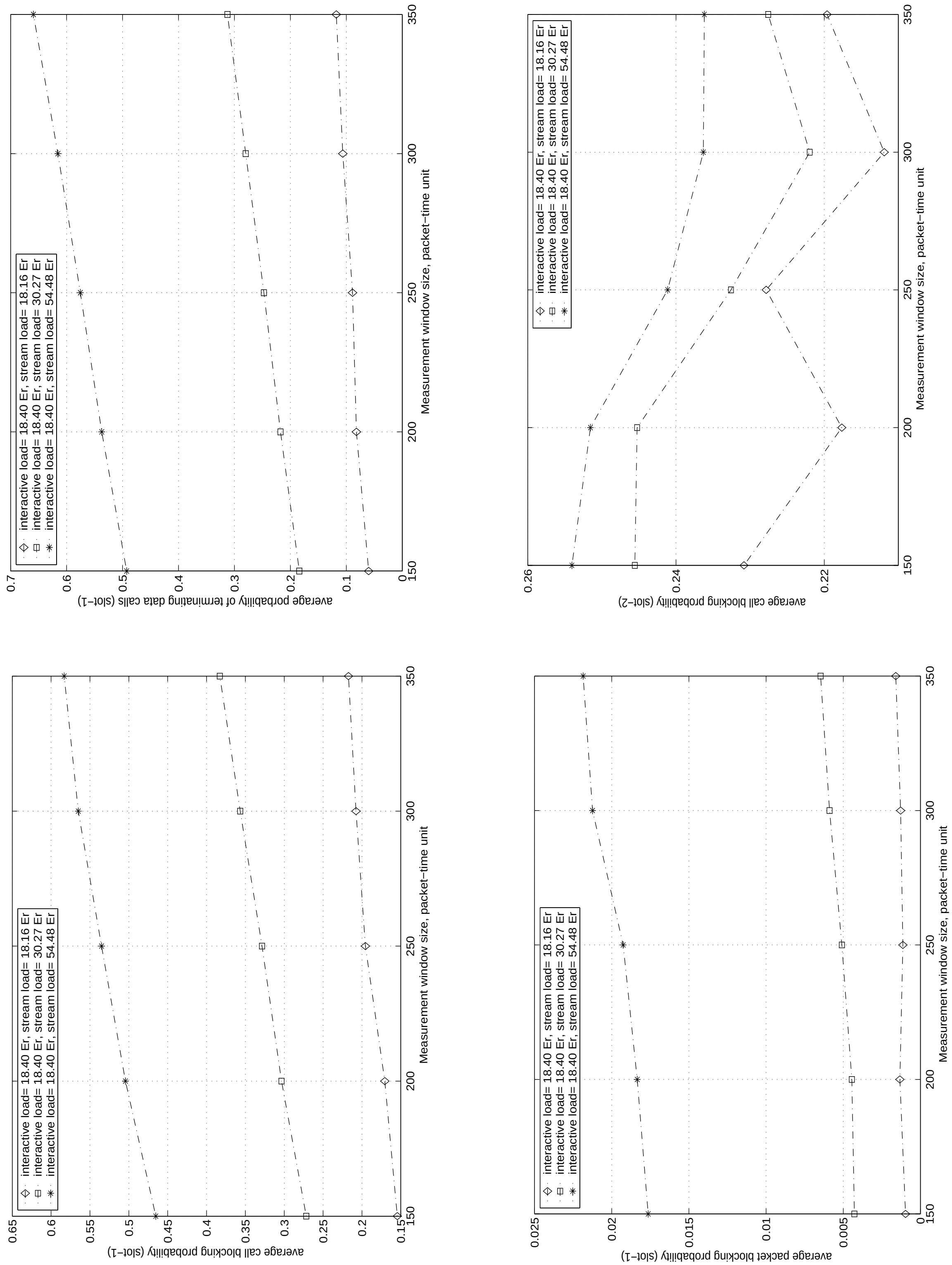


Figure 8.27: The comparison of call blocking, packet blocking and call termination probabilities versus the measurement window size under policy-I, where buffer capacity = 80, $q = 90$

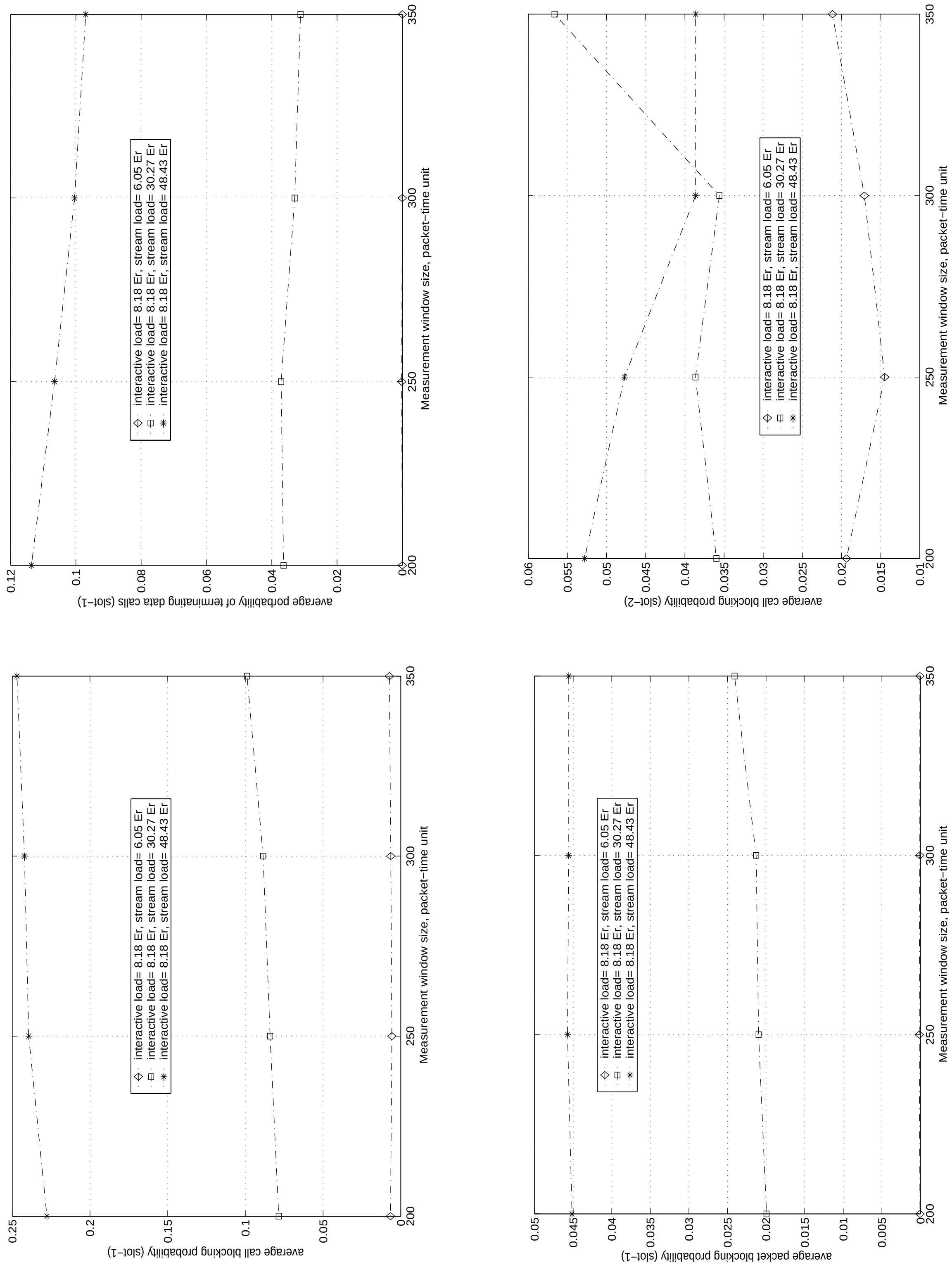


Figure 8.28: The comparison of call blocking, packet blocking and call termination probabilities versus the measurement window size under policy-II, where buffer capacity = 80, $q = 90$

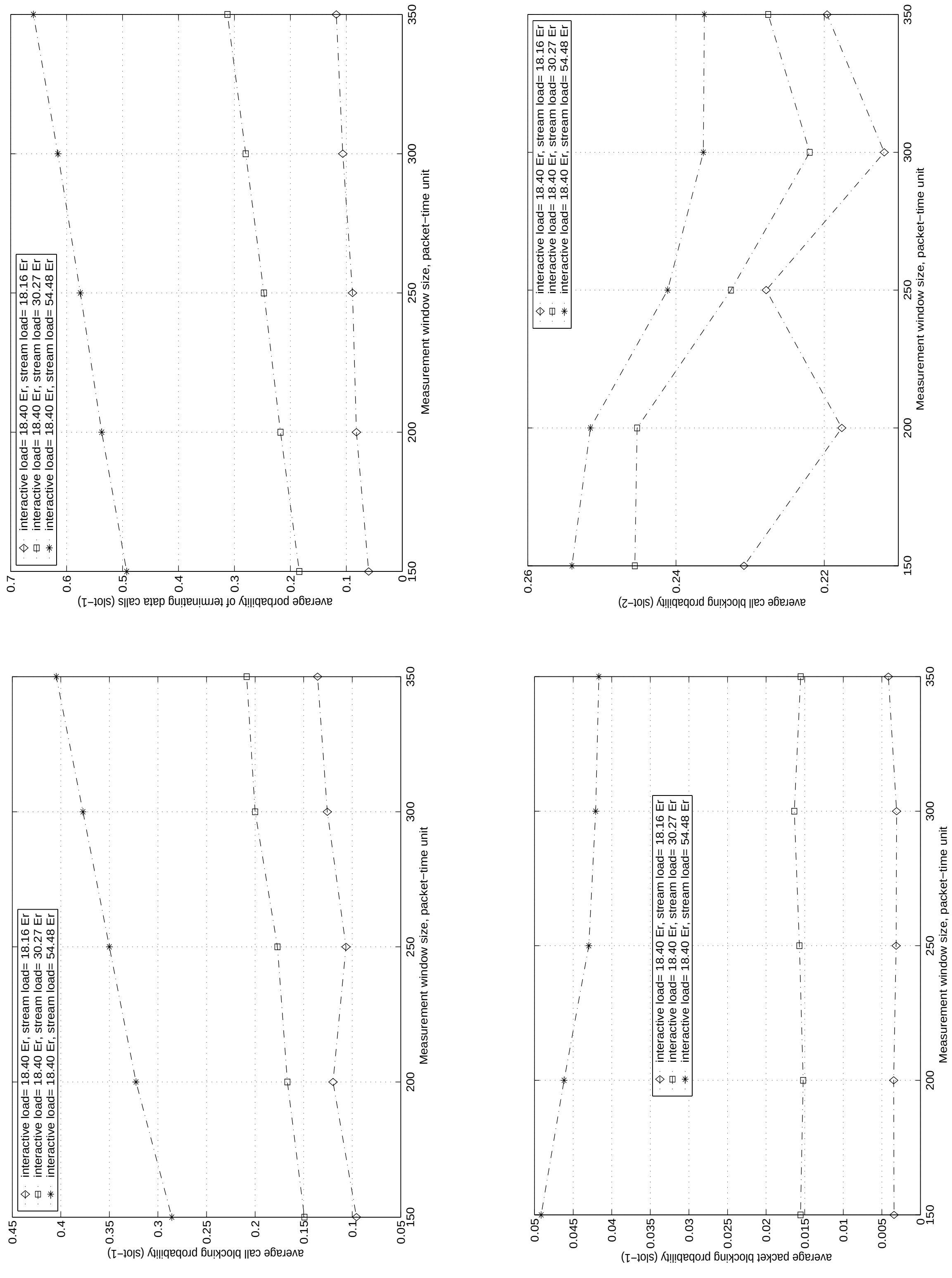


Figure 8.29: The comparison of call blocking, packet blocking and call termination probabilities versus the measurement window size under policy-III, where buffer capacity = 80, $q = 90$

a huge number of packet should go simultaneously on downlink channel. Hence, the likelihood of unsuccessful transmission on the downlink channel will increase.

More, since policy-II depends on the event of buffer occupation at the end of each measurement window (W is fixed), the probability of having the larger buffer being congested at the end of fixed measurement window gets larger. Figure 8.36 shows that the network operating under policy-2 is more sensitive to the base station buffer size than under policy-1. We notice a 70% increase in the call blocking for high stream traffic load. This happens because when we allow more packets to wait in the buffer, the probability of congested buffer will be higher as shown in Fig. 8.39 which shall result in more termination of new calls as we can see it clearly in Fig. 8.36. Therefore, these two consequences (i.e. congested buffer and more terminated calls) will contribute to higher call blocking.

On the other hand, we observe a drop (about 4%) in the call blocking probability of the interactive calls under policy-I and about 23% under policy-II. Yet, it is obvious the advantage of having larger buffer size when we consider the packet blocking probability as shown in Fig. 8.35. It shows a 39% drop in packet blocking probability when the buffer size is 100 compared to the case of just 60. This gain, however, will lead to two drawbacks. First, more packets shall reside in the buffer which means longer delay as shown in Fig. 8.38. Also, the call request shall wait longer time to get admission to the network as shown in Fig. 8.38. Second, more calls will be terminated as exploited in Fig. 8.35 as a consequence of having large number of packets awaiting in the buffer.

Considering the throughput of the network under the three policies, we can draw the following observations. First, policy-I and Policy-III do not show sensitivity for the buffer size for this performance measure. Yet, policy-II proves to be very sensitive especially under heavy stream traffic load as shown in Fig. 8.39. Again, this phenomenon is due to high call termination probability as explained above.

Figure 8.40 shows an interesting result that is the probability of non-congested

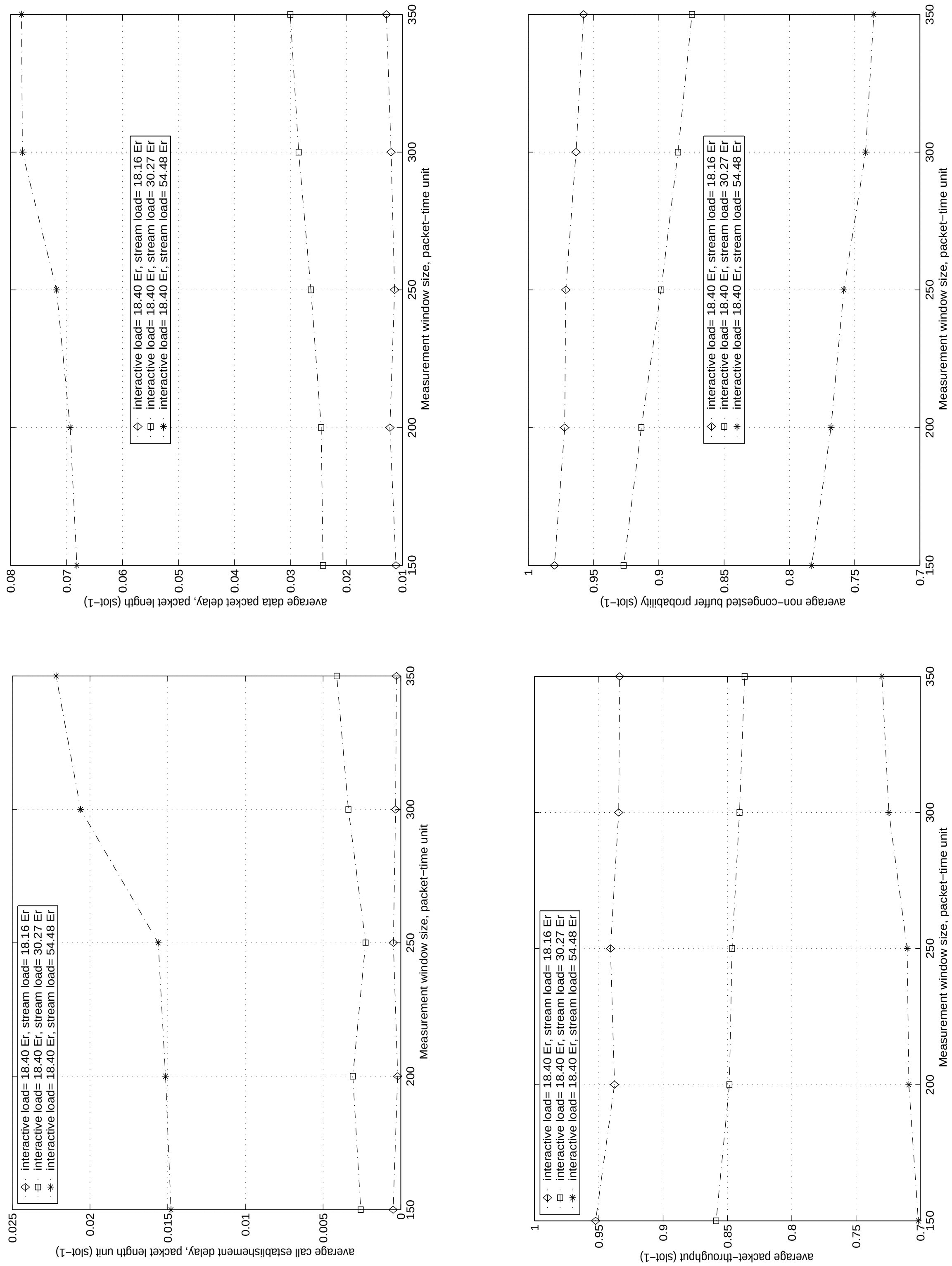


Figure 8.30: The comparison of throughput, call establishment delay, packet delay and non-congested buffer probability versus the measurement window size under policy-I, where buffer capacity = 80, $q = 90$

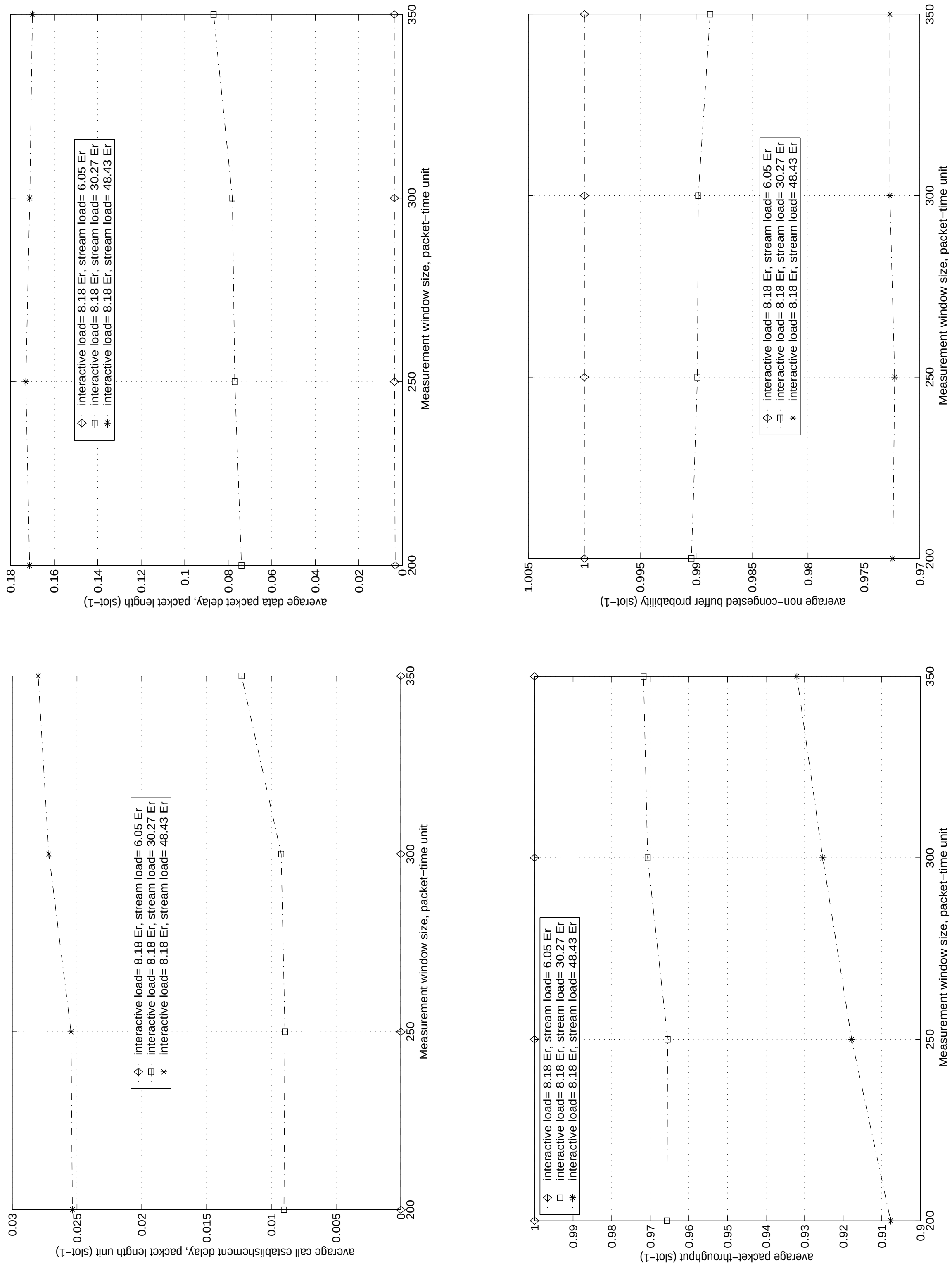


Figure 8.31: The comparison of throughput, call establishment delay, packet delay and non-congested buffer probability versus the measurement window size under policy-II, where buffer capacity = 80, $q = 90$

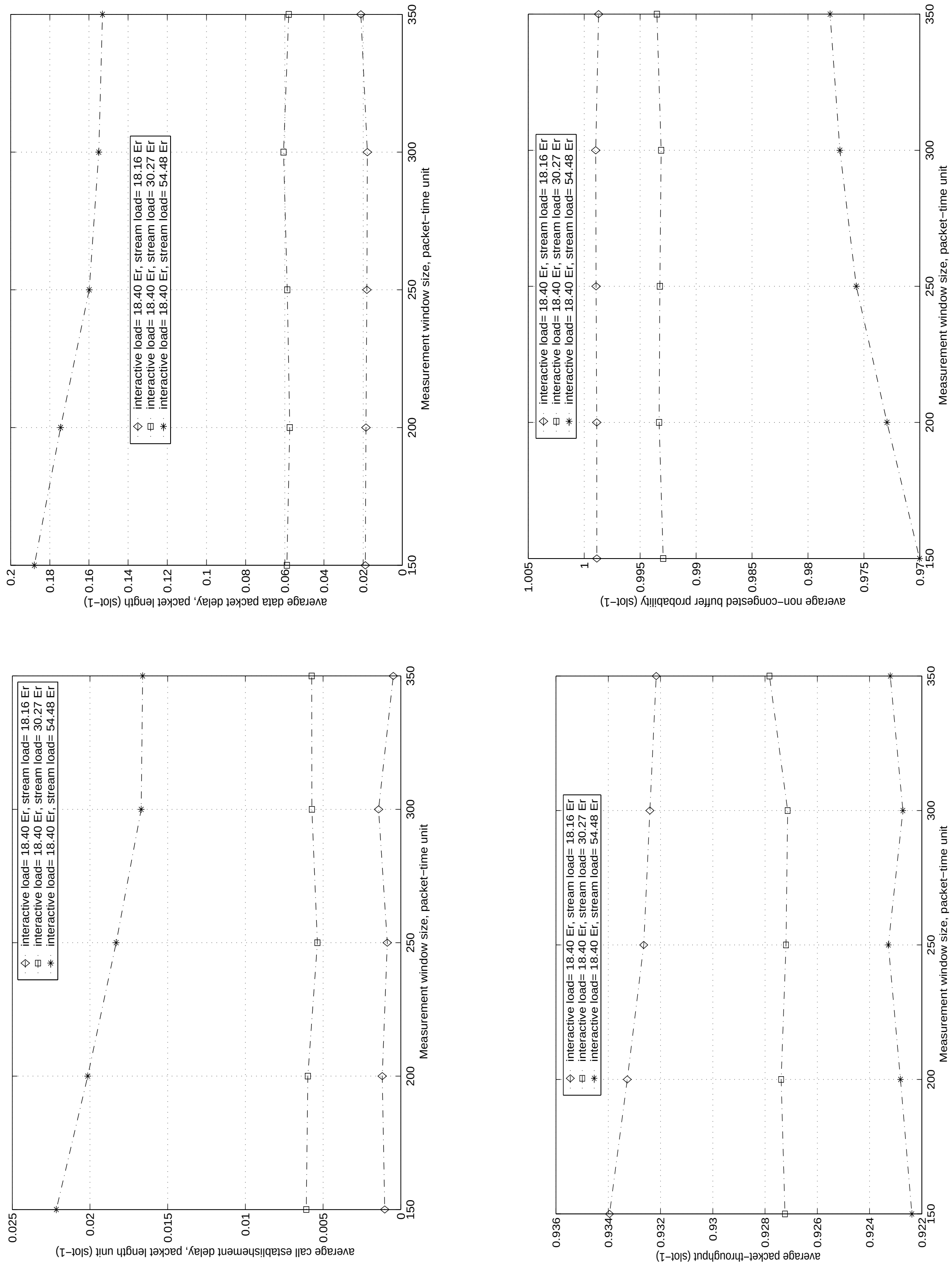


Figure 8.32: The comparison of throughput, call establishment delay, packet delay and non-congested buffer probability versus the measurement window size under policy-III, where buffer capacity = 80, $q = 90$

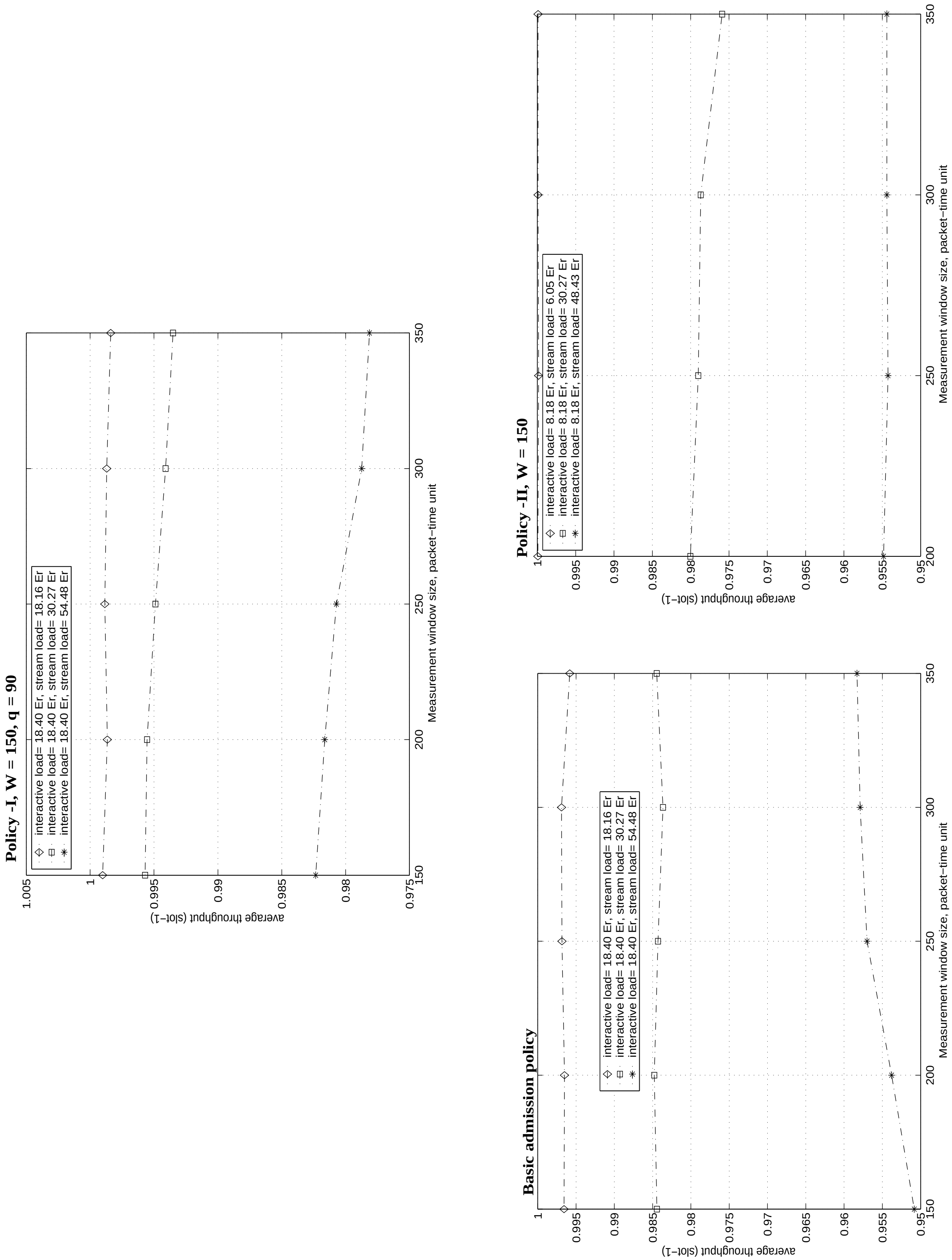


Figure 8.33: The comparison of versus the measurement window size under all proposed policies, where $B = 80$ packets, $q = 90$

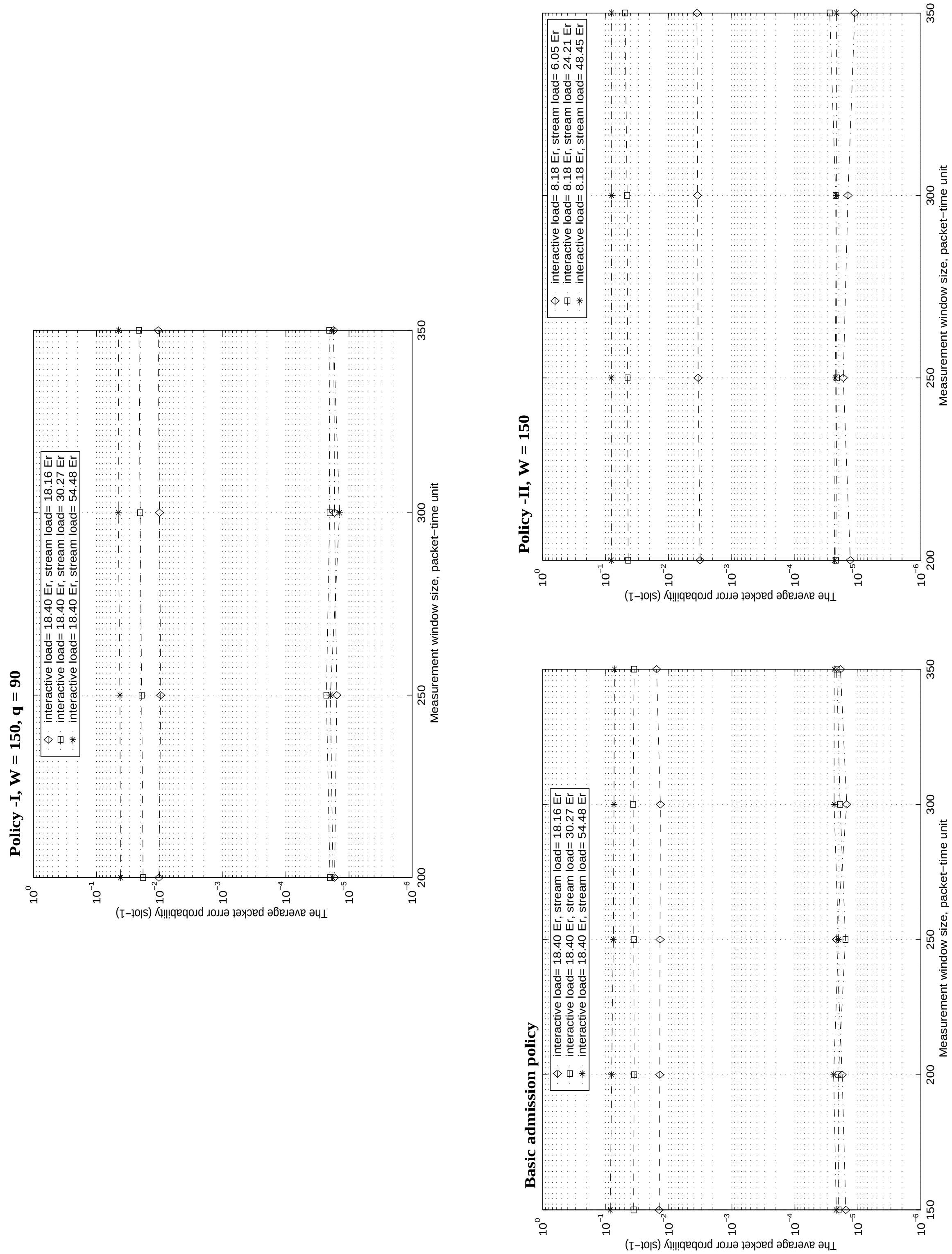


Figure 8.34: The comparison of simulation results versus the measurement window size under all proposed policies, where $B = 80$ packets, $q = 90$

under policy-III, but it has been evaluated as defined in policy-I. It is clear that under policy-III the network can not guarantee this stringent QoS.

Figure 8.38 shows the probability that the base station buffer is not congested as defined in policy-I. We notice that this performance measure is guaranteed for low and medium traffic loads regardless of the buffer size. In other words, the proposed algorithm is effective in maintaining the 90% target of non-congested buffer probability. Furthermore, Fig. 8.38 and Fig. 8.40 show a little drop in the overall packet throughput for a larger buffer size base station. This result can be attributed to the fact that the larger the number of packets in the buffer awaiting for simultaneous transmission via downlink, the larger is the loss in the number of successfully received packets by the end users.

From Fig. 8.41, the larger the base station buffer the better is the packet throughput and eventually the lower is the overall packet losses. Moreover, Policy-I outperforms the other two policies regarding this performance measure.

The design of the network under consideration gives higher priority for delay-sensitive applications. Hence, these applications do not experience any delay except the end-to-end propagation delay and processing delay. Yet, it is crucial to evaluate the overall packet losses for these applications. Figures 8.13-8.15 illustrate the packet losses on the down link where the three policies achieve the same level of quality. This is because these policies are designed to enhance the performance of stream traffic when they are integrated with interactive traffic. Furthermore, Fig. 8.42 depicts the average packet error probability for the traffic using slot-I. It is obvious that the uplink performance is much better than downlink since multiuser decorrelator receiver has been used for uplink while a single user receiver is used for downlink traffic. Since a retransmission mechanism is adopted for downlink traffic the high packet error probability play a little role in packet losses. The major factors for packet losses for stream users are the uplink error probability and the packet drop due to buffer overflow.

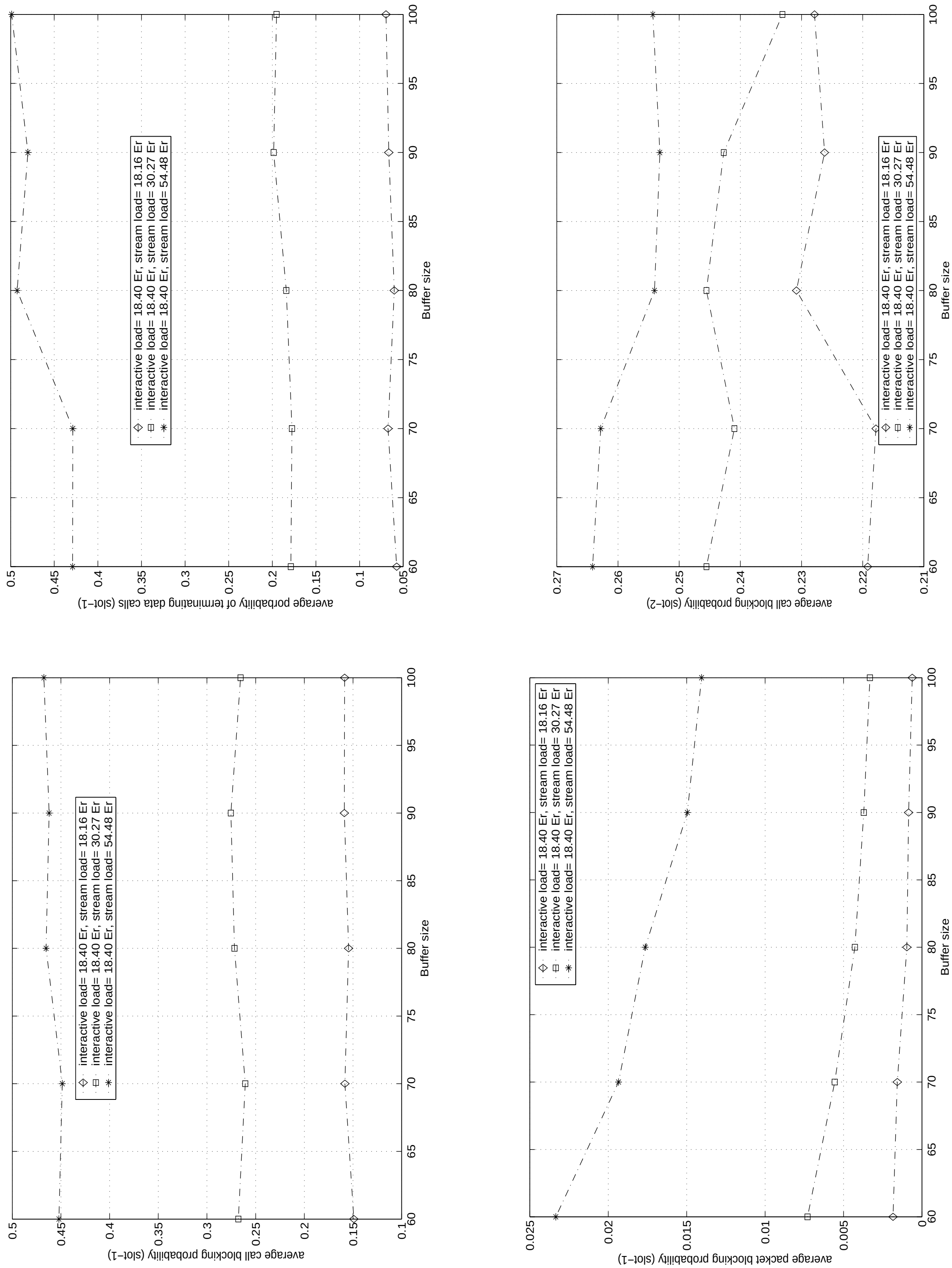


Figure 8.35: The comparison of call blocking, packet blocking and call termination probabilities versus the buffer size at the base station under policy-I, where $W = 150$ packets, $q = 90$

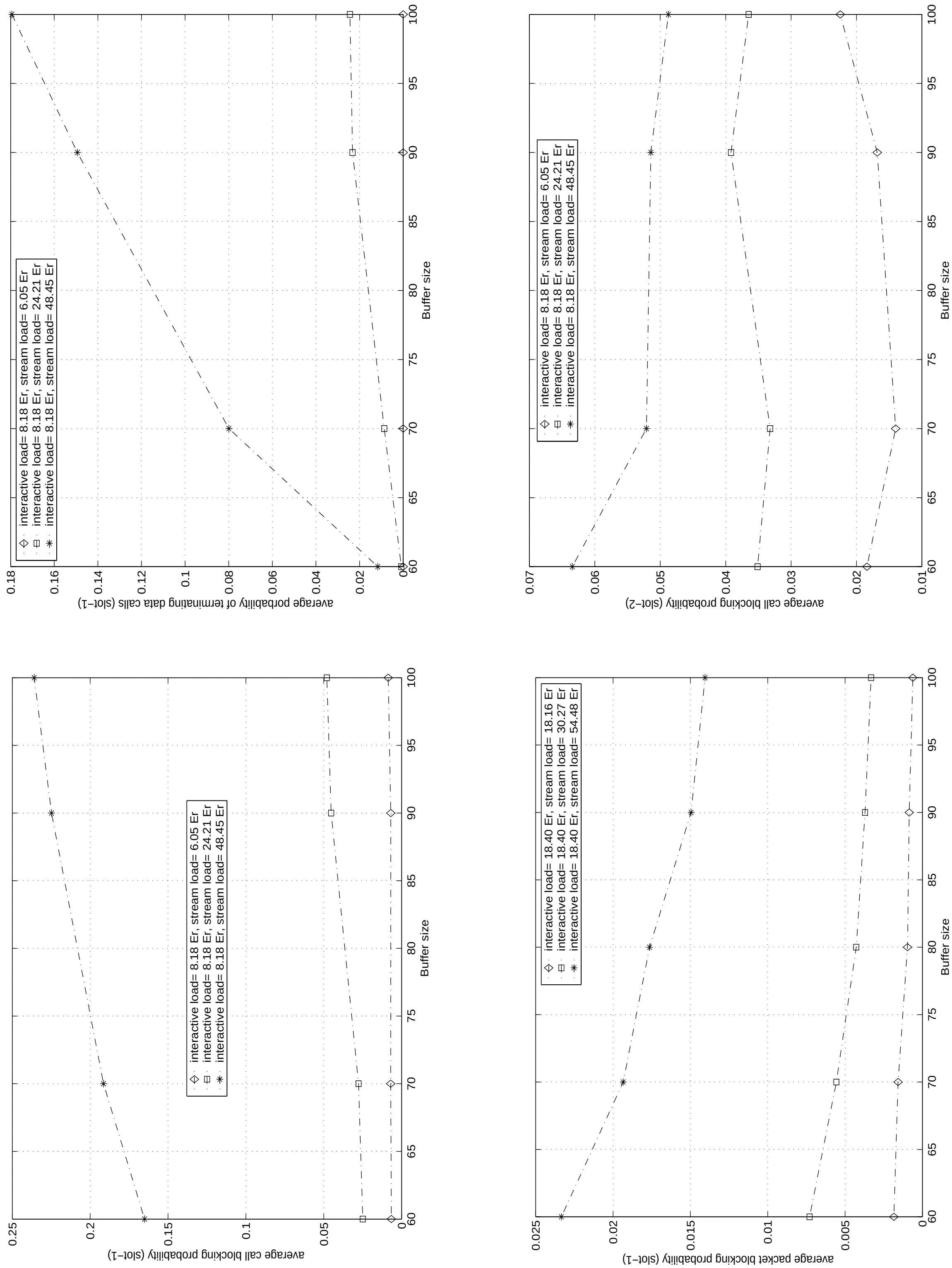


Figure 8.36: The comparison of call blocking, packet blocking and call termination probabilities versus the buffer size at the base station under policy-II, where $W = 150$ packets

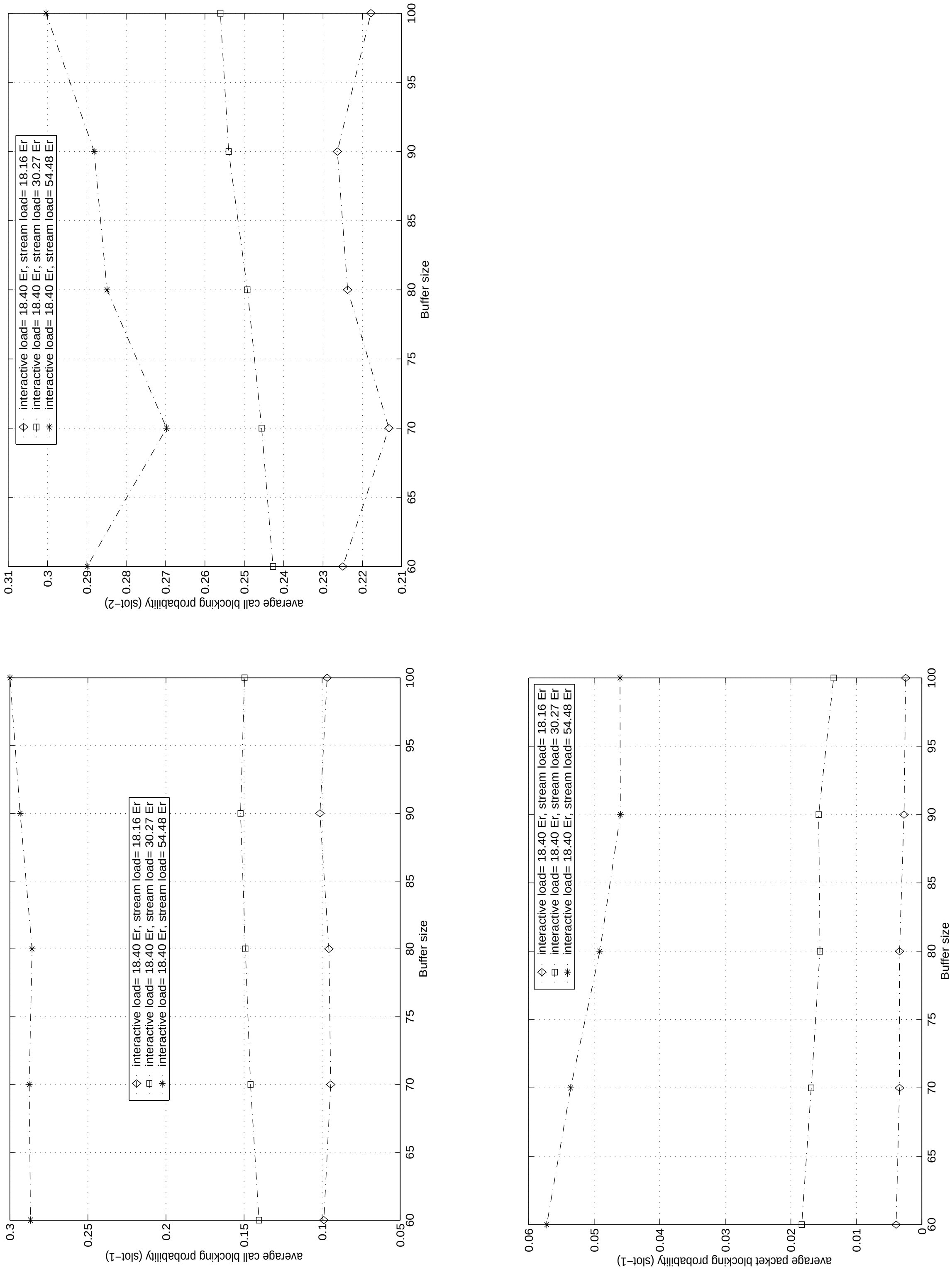


Figure 8.37: The comparison of call blocking, packet blocking and call termination probabilities versus the buffer size at the base station under policy-III, where $W = 150$ packets

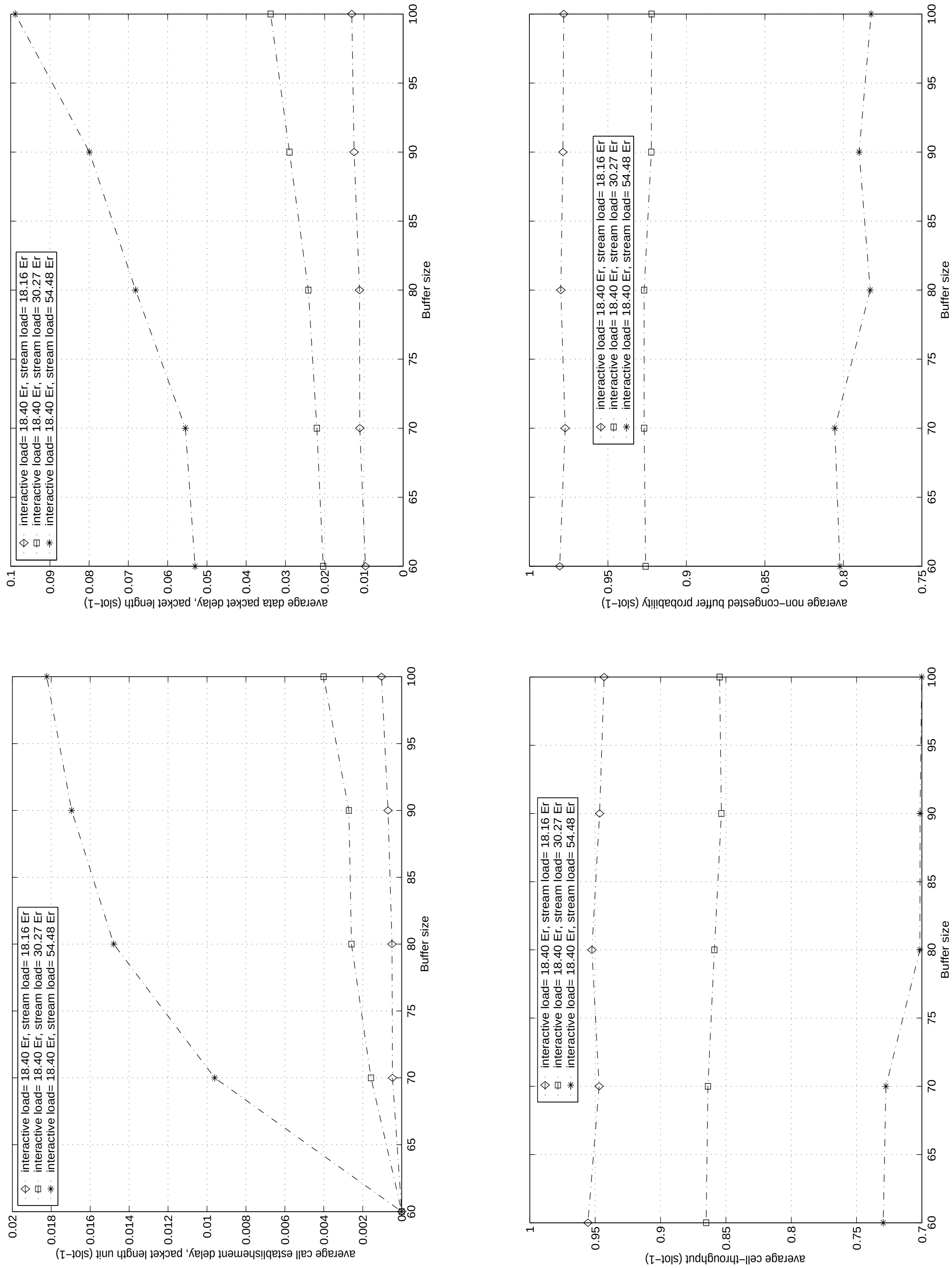


Figure 8.38: The comparison of throughput, call establishment delay, packet delay and non-congested buffer probability versus the buffer size at the base station under policy-I, where $W = 150$ packets, $q = 90$

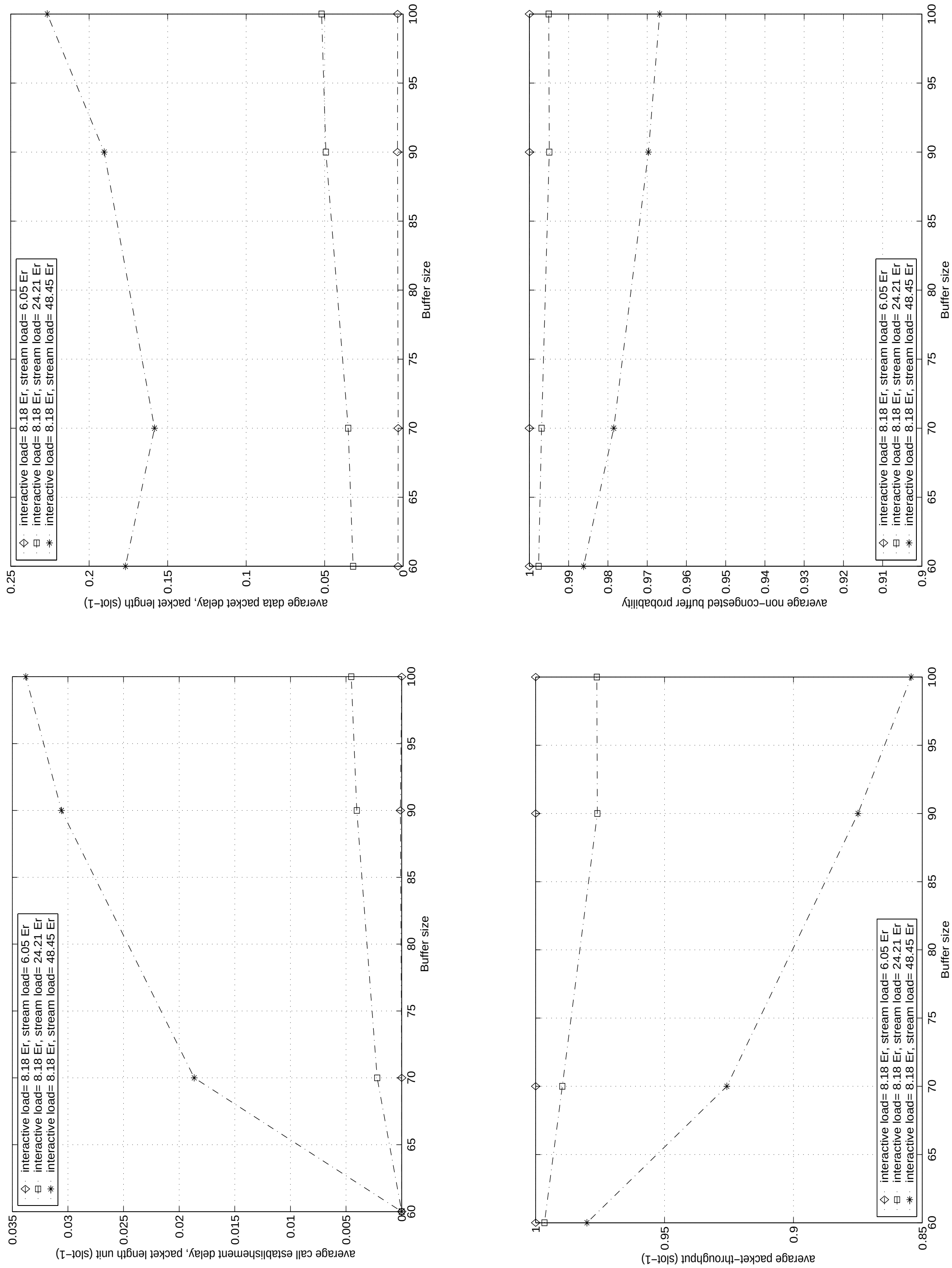


Figure 8.39: The comparison of throughput, call establishment delay, packet delay and non-congested buffer probability versus the buffer size at the base station under policy-II, where $W = 150$ packets

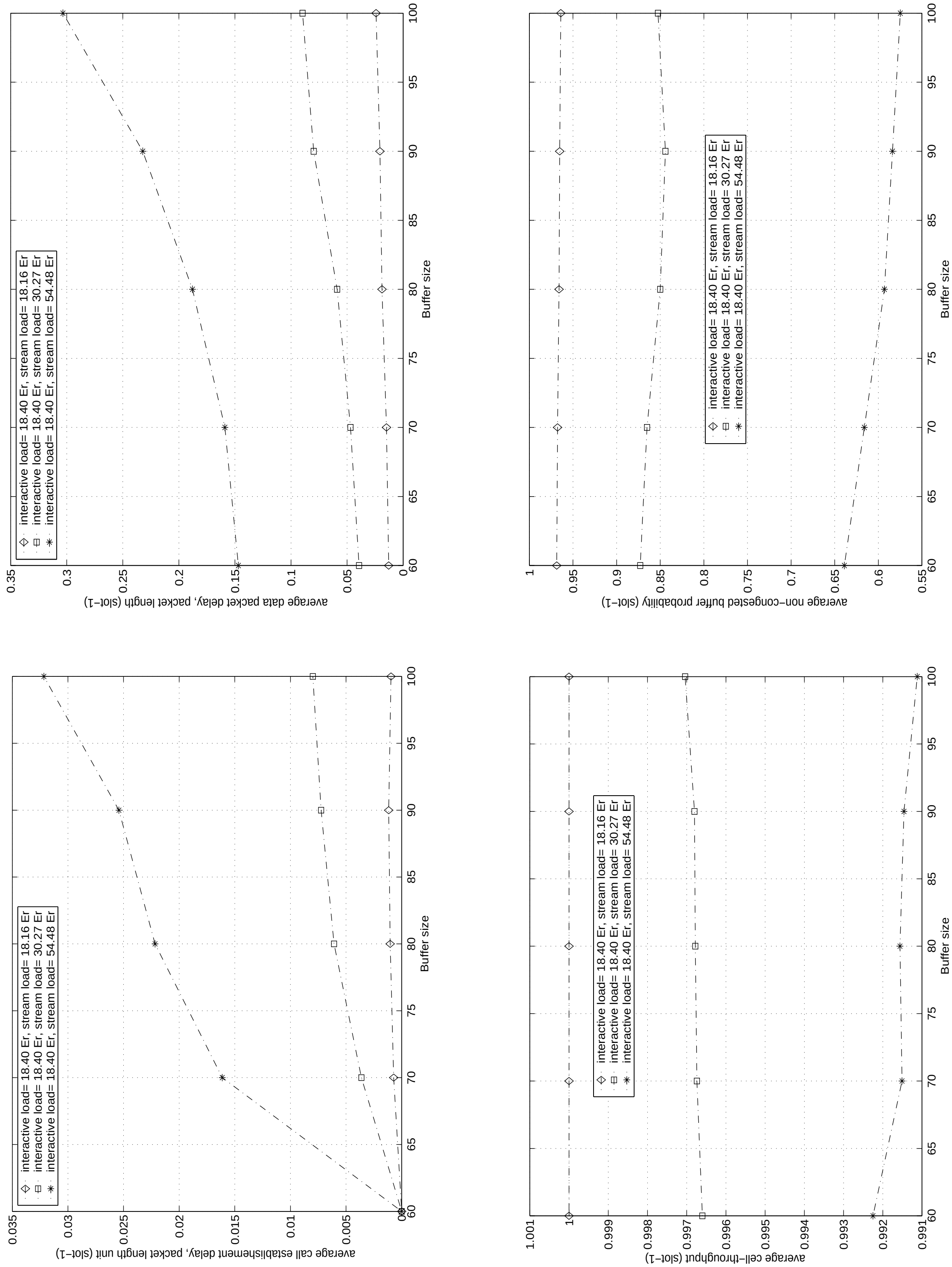


Figure 8.40: The comparison of cell-throughput, call establishment delay, packet delay and non-congested buffer probability results versus the buffer size at the base station under policy-III, where $W = 150$ packets

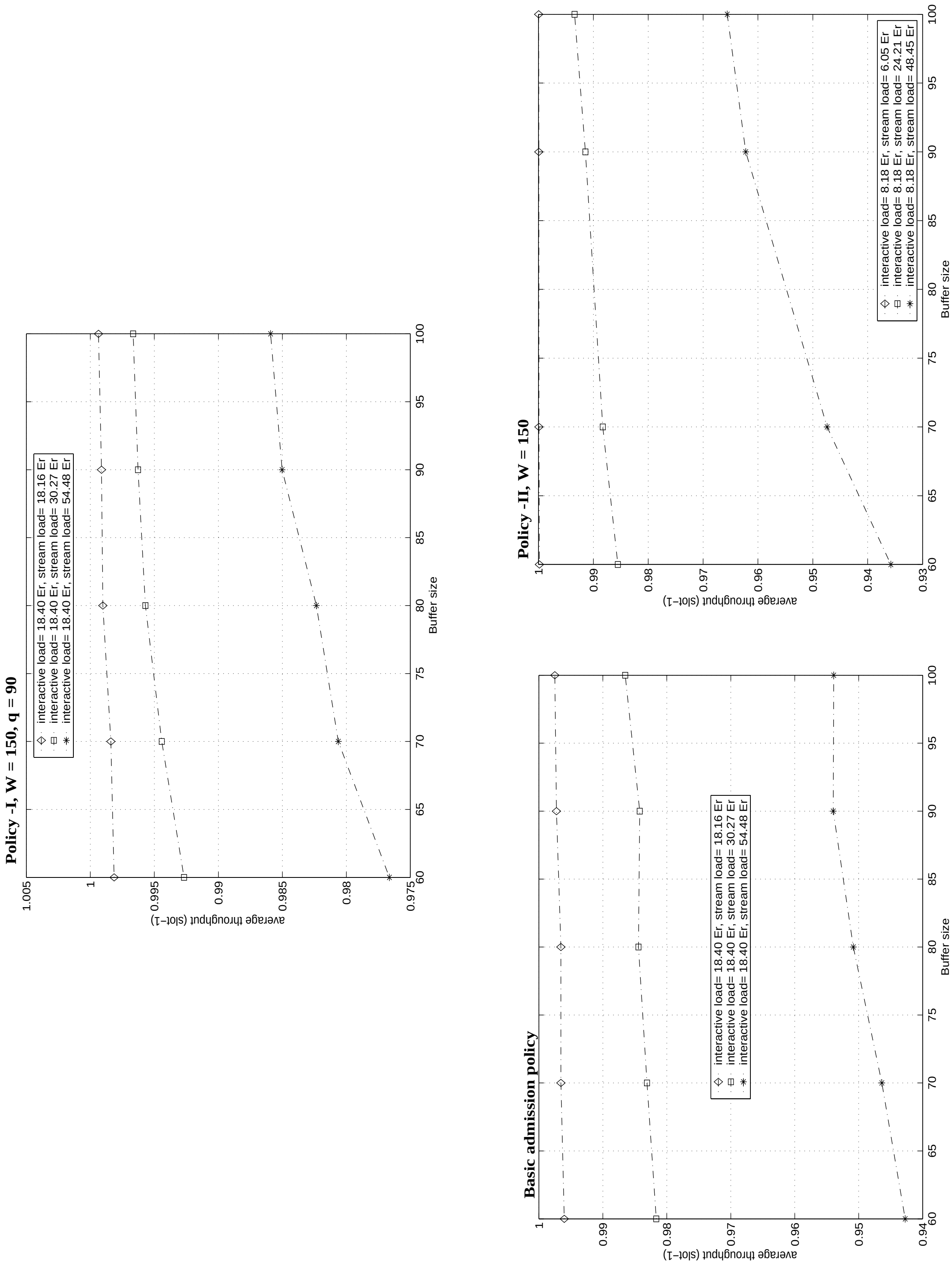


Figure 8.41: The comparison of simulation results of average throughput, under all proposed policies versus the buffer size at the base station, where $W = 150$ packets, $q = 90$

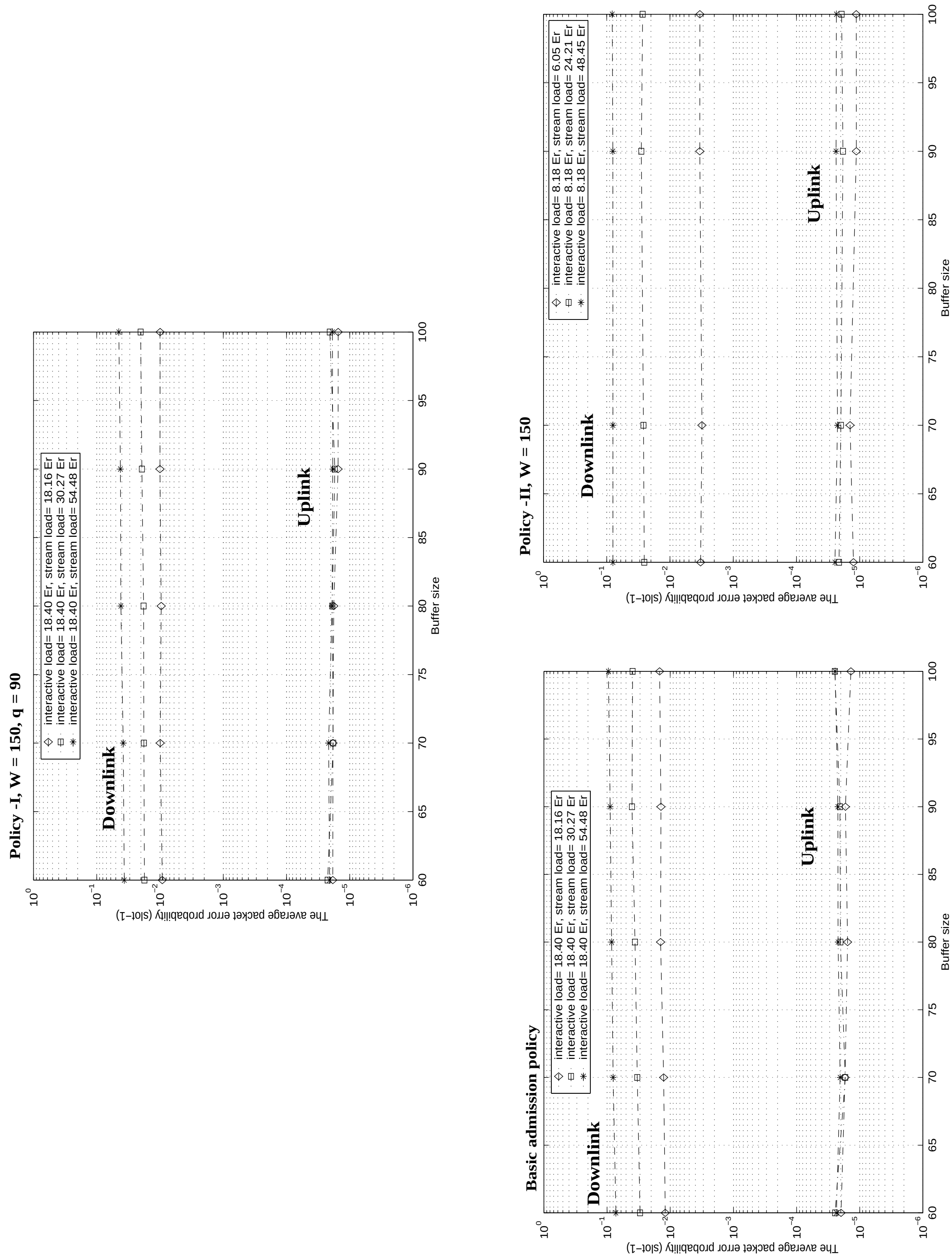


Figure 8.42: The comparison of simulation results of packet error rate of the traffic in slot-1, under all proposed policies versus the buffer size at the base station, where $W = 150$ packets, $q = 90$

Tables 8.3-8.6 show coma prison summaries for all simulated admission/congestion policies where policy-III is used as a reference. The '+' means a gain from the network performance standpoint, while '-' means a loss from the network performance standpoint. For example, table 8.3 shows a five-fold increment in call blocking under policy-I compared to policy-III, while policy-II enhance the call blocking by only 77%. These enhancements in the call blocking are actually losses from the network standpoint.

On the other hand, packet blocking probability has drop ed by 67.4% when we apply policy-I and by only 10% when policy-II is applied compared to policy-III as shown in table 8.5. Of course, this drop in packet blocking is an improvement in the network performance. Further, from table 8.5 and table 8.6 we can observe an interesting result that is at high traffic load, policy-I outperforms other policies for only 15% increment in call blocking and at the same time maintaining its stringent QoS.

8.7 Conclusions

Chapter 7 is devoted completely to analytically analyze the proposed admission / congestion policy on "pure" CDMA and hybrid MC-CDMA/TDMA access protocols. Here, in this chapter, several scenarios of the proposed admission / congestion policy to accompany the hybrid MC-CDMA/TDMA access protocol have been evaluated using Monte-Carlo simulation. In spite of the fact that each method has different approach, both methods led independently to the same qualitative conclusions. Also, the obtained results show good agreement quantitatively, yet there are some discrepancies due to the assumptions used in each method. For example, we notice that the results obtained for some performance measures under the analytical method show sudden drop beyond certain traffic load. In contrast, the ones obtained from simulation are more smoother. This is attributed to the aggressive adaptation policy applied in the analytical analysis. In conclusion, these results indicate that both methods are reliable in evaluating the network performance measures under the proposed policies.

Table 8.3: Comparison of the network performance under the three proposed policies, $\rho_{Int} = 6.13$ Erlang, $\rho_S = 30.27$ Erlang

Performance Measure	Policy-I ($q = 90$)	Policy-I ($q = 30$)	Policy-II	Policy-III	Gain/Loss (Policy-I)	Gain/Loss (Policy-II)
Call Blocking (stream)	0.2101	0.2106	0.0630	0.03563	489% (-)	77% (-)
Call Blocking (interactive)	0.0078	0.0092	0.0207	0.0095	17.9% (+)	118% (+)
Packet Blocking	0.00662	0.00665	0.0192	0.02469	73% (+)	22% (+)
Packet Delay	0.0323	0.0326	0.0719	0.0893	63.8% (+)	19.5% (+)
Congested Buffer	0.8962	0.9202	0.9906	0.9879	okay	okay
Throughput	0.8137	0.8148	0.9606	1.0000	18.6% (-)	3.9% (-)
Packet Throughput	0.9934	0.9933	0.9807	0.9953	1.85% (+)	0.5% (+)
Call Termination	0.2556	0.2567	0.03862	0	(-)	(-)

Table 8.4: Comparison of the network performance under the three proposed policies, $\rho_{Int} = 6.13$ Erlang, $\rho_S = 66.56$ Erlang

Performance Measure	Policy-I ($q = 90$)	Policy-I ($q = 30$)	Policy-II	Policy-III	Gain/Loss (Policy-I)	Gain/Loss (Policy-II)
Call Blocking (stream)	0.4847	0.4843	0.3444	0.2872	68.9% (-)	19.9% (-)
Call Blocking (interactive)	0.0163	0.0207	0.0237	0.0307	46.9% (+)	22.8% (+)
Packet Blocking	0.0358	0.0359	0.06473	0.0810	55.8% (+)	16.9% (+)
Packet Delay	0.1336	0.1340	0.2700	0.3663	63.5% (+)	26.3% (+)
Congested Buffer	0.6272	0.6987	0.9553	0.9413	(-)	(-)
Throughput	0.6017	0.6038	0.8360	1.0000	39.8% (-)	16.4% (-)
Packet Throughput	0.9642	0.9641	0.9353	0.9189	4.9% (+)	1.77% (+)
Call Termination	0.7827	0.7717	0.2055	0	(-)	(-)

Table 8.5: Comparison of the network performance under the three proposed policies, $\rho_{Int} = 67.48$ Erlang, $\rho_S = 30.27$ Erlang

Performance Measure	Policy-I ($q = 90$)	Policy-I ($q = 30$)	Policy-II	Policy-III	Gain/Loss (Policy-I)	Gain/Loss (Policy-II)
Call Blocking (stream)	0.4271	0.4197	0.3760	0.3715	15% (-)	1.2% (-)
Call Blocking (interactive)	0.7002	0.6985	0.7031	0.7003	0%	0%
Packet Blocking	0.0017	0.0016	0.0046	0.0051	67.4% (+)	10.3% (+)
Packet Delay	0.0141	0.0140	0.0234	0.0248	43.1% (+)	5.8% (+)
Congested Buffer	0.9694	0.9774	0.9983	0.9982	okay	okay
Throughput	0.92753	0.9292	0.99616	1.0000	7.2% (-)	0.38% (-)
Packet Throughput	0.9983	0.9984	0.9954	0.9949	0.35% (+)	% (+)
Call Termination	0.0908	0.0910	0.0062	0	(-)	(-)

Table 8.6: Comparison of the network performance under the three proposed policies,$\rho_{Int} = 67.48$ Erlang, $\rho_S = 66.59$ Erlang

Performance Measure	Policy-I ($q = 90$)	Policy-I ($q = 30$)	Policy-II	Policy-III	Gain/Loss (Policy-I)	Gain/Loss (Policy-II)
Call Blocking (stream)	0.6382	0.6323	0.5604	0.5442	17.3% (-)	3.0% (-)
Call Blocking (interactive)	0.7148	0.7097	0.7191	0.7200	0%	0%
Packet Blocking	0.0083	0.0079	0.0209	0.0265	68.5% (+)	21.0% (+)
Packet Delay	0.0369	0.03567	0.0775	0.0958	61.43% (+)	19.1% (+)
Congested Buffer	0.8914	0.9218	0.9930	0.9871	okay	okay
Throughput	0.7895	0.79741	0.9546	1.0000	21% (-)	4.5% (-)
Packet Throughput	0.9916	0.9921	0.9791	0.9735	1.86% (+)	0.57% (+)
Call Termination	0.3095	0.2888	0.0465	0	(-)	(-)

Chapter 9

Conclusion and Future Work

9.1 Conclusions

We have established means to improve the performance of DS-CDMA integrated networks. These proposed protocols especially the Measurement-based admission / congestion policies can easily be applied to other networks.

Motivated by the need for integrating heterogeneous services on wireless networks as well as the backbone wireline networks, two novel traffic flow control approaches accompanied a hybrid TDMA/MC-CDMA system utilizing multiuser detection have been proposed and studied. One approach deterministically controls the flow of traffic into the TDMA slots [94], while the other statistically controls the flow of traffic depending on the instantaneous changes in the traffic load. The essence of our work is to introduce an interaction between the physical layer and higher layers, enabling more practical utilization of multiuser detection and supporting services with different QoS parameters.

To facilitate this interaction, a traffic classification is introduced based on the fact that each category has its own required QoS as well as its own traffic and transmission characteristics. Hence, each category of traffic should be treated differently. In the literature, just few previous studies in DS-CDMA networks did take into consideration the relationship between the physical layer design and traffic flow traversing this physical layer.

More, a unified approach is introduced that is the strategy of subdividing the high rate users into low-rate streams and then applying the CDMA technique. This strategy (called Multi-Code CDMA (MC-CDMA)) has very interesting features that the other systems lack. It can easily integrate traffic of different transmission rates in a unified structure where all the transmissions over the radio channel occupying the same bandwidth and having the same processing gain [91]. The MC-CDMA technique is general and it can be implemented on top of any switching protocols such as circuit switching, packet switching, ATM, etc. For instance, consider the ATM technology, the accommodation of VBR transmission through implementation of MC-CDMA along with the proposed traffic control policies guarantees relative simplicity and speed compared to ATM accommodation mechanism of VBR users. However, there is nothing in our new technique that could forbid the added benefits of ATM VBR technique on top where the proposed techniques in this paper could be adopted in the ATM adaptation layer (AAL).

The analytical analysis has shown the outstanding performance of the proposed access protocol compared with the conventional DS-CDMA network. This improvement in the performance is not only due the utilization of Multiuser detector but also is attributed to the novel flow traffic control schemes that have introduced an interaction between the physical layer and higher network layers and consequently balance the traffic load on each detection algorithm and makes the implementation of the decorrelator receiver more practical. The simplicity and effectiveness of the proposed system make it very suitable for future 3G CDMA where heterogeneous traffic sources are expected.

Further, a Measurement-Based admission/congestion policy has been proposed and studied. The proposed admission/congestion policy has two phases. The Phase-I (admission) manages the admission of new calls such that there are enough resources (servers) on the uplink to serve these new users. The proposed traffic flow control mentioned above imposes the following restriction. Stream users should be admitted into the system, if the available servers can accommodate the variable rate components (excess traffic) of active stream users concurrently with the active interactive users. This means that the admission policy grants higher priority to interactive users calls which is very practical for delay-sensitive application such as voice, video calls. On the the hand, we can assume the same level of accessibility for both traffic categories providing that the stream users should restrict their transmission rates to the average rates if all servers during slot-II would be busy serving the interactive users. This modification would lead to a larger delay especially for high variable rate users.

Secondly, the Phase-II (congestion) of the proposed traffic control policy manages the packet flow into the base station such that it prevents any possible congestion in the buffer. This objective is achieved by interrelating the End-to-End network parameters such as up/down link channels, number of receivers at the base station, buffer limitation at base station, traffic burstiness, etc. in one *Queueing Problem*. Subsequently, the instantaneous buffer overflow at the base station is used to predict the likelihood of buffer congestion over a period of window of transmitted packets. This new queuing model has very interesting application where it can be fed by real data and used 'off-line' to predict the network performance.

The proposed congestion policy is reliable, simple and adaptive and can be applied through different scenarios. When this policy indicates that the current offered traffic load should be reduced, then the network operator can achieve this objective through several ways depending on his contracts with the network users. For example, terminating a call would not be an option for users who pay high fees, but it could be an option for users who just pay for available bandwidth and so on.

The practicality and several performance measures of the new system have been analyzed analytically under a wide range of expected traffic characteristics (bit rate, transmission activities, etc.) for the future wireless networks. The adaptive admission/congestion control policy based on Window-Measurement has effectively maintained the required QoS, particularly the blocking probability, call establishment delay and cell error rate. It is worth to emphasis here that our proposed admission/congestion control polices apply preventive admission control as well as reactive congestion control.

The complexity of the wireless networks under consideration has put some restrictions on the analytical analysis. Hence, the final part of this work was to evaluate the performance of the proposed protocols via simulation where these protocols shall be tested under more realistic conditions and system parameters. We use exact Monte-Carlo simulation to simulate the multimedia integrated CDMA networks where heterogeneous traffic users are multiplexed into a simple two slots TDMA frames. In this simulation, our emphasis has been on the network performance issues such as packet delay, packet delay jitter, packet losses, call blocking, etc. Therefore, the only system parameters that have been simulated and randomly generated throughout the simulation program are these parameters which have direct relation to the network performance. For instance, packet error performance will not be simulated. However, the Monte-Carlo simulation is supplemented by analytical computation of packet error probability for uplink stream packets received by decorrelator receiver at the base station (BS). On the other hand, the packet error probability for the uplink slot-II users as well as downlink users from both slots are supplemented from computational analysis.

In addition, because of the relative simplicity of simulation compared with analytical analysis, the original proposed admission/congestion policy has been enhanced and extended during the course of simulation. therefor, we ended with three Measurement-Based admission/congestion policies. Each can support certain level of QoS requirements (i.e. delay call blocking, packet blocking, etc.).

Tables 8.3-8.6 show a brief summary of the network performance under these policies. These tables exploit the advantages and disadvantages of the proposed admission/congestion policies. It is worth to note that some performance measures can not be compared such as probability of non-congested buffer since each has been evaluated under different scale. The only disadvantage of policy-I is the high call blocking probability compared to the other two polices especially under low stream traffic load. Nevertheless, we find policy-I outperforms other polices when we consider other performance measures such as packet delay, packet throughput, etc.). Therefore, the question is which policy should the network operator choose? Of course, the answer is a matter of trade off between the QoS requirements.

9.2 Future Work

The research area of multiple access techniques along with admission/congestion control policies is very wide. The work in this thesis just covers a small part of this very important state of the art. Hence, this work can be extended easily. The following topics are just a few.

1. Investigating the end-to-end QoS performance of the CDMA network under the proposed admission/congestion policy taking into consideration different queueing schemes other than priority queue at the user side, such as custom queue, Fair weighted queue, etc.

2. Building a complete simulation package that takes into consideration all system's parameters such as multipath channel, error correction coding, user mobility, etc.

3. Investigating more the 2^{nd} flow control approach accompanied with admission/congestion control policy as well as Kalman filtering traffic estimator or any other estimation technique.

4. Considering the issue of buffering delay-sensitive traffic, but for controlled delay constraint at the BS as well as at the user side.

5. Investigating more the relation between the measurement window size, buffer size at BS and at the end user and the statistical models of the traffic using the network.

Bibliography

[1] "Special Issue on Wireless ATM," *IEEE Personal Commun.*, August 1996.

[2] "Special Issue on Spread Spectrum," *IEEE Tran. on Commun.*, vol. Com-30, no. 5, May 1982.

[3] "Special Issue on Code Division Multiple Access Networks II," *IEEE J. Select. Areas Commun.*, vol. SAC-12, no. 5, June 1994.

[4] "Special Issue on Code Division Multiple Access Networks III," *IEEE J. Select. Areas Commun.*, vol. SAC-14, no. 8, October 1996.

[5] "Special Issue on Third Generation land Mobile Communication Systems," *IEICE Tran. on Commun.*, vol. E81-B, no. 7, July 1998.

[6] F. Adachi, M. Sawahashi, and H. Suda, "Wideband DS-CDMA for Next Generation Mobile Communications Systems,"*IEEE Communications Magazine*, pp. 56-69, September 1998.

[7] C. P. Tou and D. A. Roy, "On efficient Spectrum Utilization from the Standpoint of Communication Theory," IEEE Trans. Info. Theory, vol. IT-68, no. 12, Dec. 1980, pp. 1460-1465.

[8] J. G. Proakis, *Digital Communications*, 3rd ed., New York: McGraw Hill, 1995.

[9] R. Rom, and M. Sidi, Multiple Access Protocols: Performance and Analysis, New York: Springer-Verlag, 1990.

[10] D. E. Everitt, "Traffic Engineering of the Radio Interface for Cellular Mobile Networks," Proc. IEEE, vol. 82, no. 9, Sep. 1994, pp. 1371-1382.

[11] J. R. Foerster and L. B. Milstein, "Analysis of Hibrid, Coherent FDMA/CDMA Systems in Ricean Multipath Fading," IEEE Tran. Commun., vol 45, no. 1, Jan. 1997, pp. 15-18.

[12] A. K. Elhakeem, R. D. Girolamo, I. B. Bdira, and M. Talla, "Delay and Throughput Characteristics of TH, CDMA, TDMA, and Hybrid Networks for Multipath Faded Data Transmission Channels," IEEE J. Select. Areas Commun., vol. SAC-12, no. 4, May 1994, pp. 622-637.

[13] N. Wilson, R. Ganesh, K. Joseph and D. Raychaudhuri, "Packet CDMA versus Dynamic TDMA for Access Control in an Integrated Voice/Data PCN," IEEE J. Select. Areas Commun., vol. SAC-11, no. 6, Aug. 1993, pp. 870-884

[14] K. Seki, Y. Sanada, Q. Wang, S. Kato and V. K. Bhargava, "A Wireless Multimedia Networks on a Time Division Duplex CDMA/TDMA," IEICE Tran. Commun., vol. E78-B, no. 7, July 1995, pp. 1002-1015.

[15] K. Kamerlander, "Benefits of Combined TDMA/CDMA Operation for Third Generation Mobile Radio Systems," Proc. of ISSSTA '96, pp. 507-512.

[16] J. J. Spilker, JR., Digital Communications by Satellite, New Jersey: Prentice-Hall, INC., 1977.

[17] F. Borgonovo, M. Zorzi, L. Fratta, V.Trecordi, and G. Bianchi, " Capture Division Packet Access for Wireless Personal Communications", IEEE J. Select. Areas Commun., vol. SAC-14, no. 4, May 1996, pp. 609-622.

[18] M. B. Pursley, "The Role of Spread Spectrum in Packet Radio Networks", Proc. IEEE, vol. 75, no. 1, Jan. 1987, pp. 116-134.

[19] A. J. Viterbi, "The Evolution of Digital Wireless Technology from Space Exploration to Personal Communication Services", IEEE Trans. Veh. Technol., vol. 43, no. 3, Aug. 1994, pp. 638-644.

[20] M. Frullone, G. Riva, P. Grazioso, and C. Carciofi, "PRMA Performance in Cellular Environments with Self-Adaptive Channel Allocation Strategies", IEEE Trans. Veh. Technol., vol. 45, no. 4, Nov. 1996, pp. 657-665.

[21] B. C. Kim, and C. K. Un, "An efficient Wireless Voice/Data Integrated Access Algorithm in Noisy Channel Environments", IEICE Tran. Commun., vol. E79-B, no. 9, Sep. 1996, pp. 1394-1403.

[22] X. Qiu and V. O. K. Li, "Dynamic Reservation Multiple Access (DRMA): A New Multiple Access Scheme for Personal Communication Systems (PCS)", Wireless Networks 2, 1996, pp. 117-128.

[23] S. Kumar and D. R. Vaman, "An Access Protocol for Supporting Multiple Classes of Services in a Local Wireless Environment", IEEE Trans. Veh. Technol., vol. 45, no. 2, May 1996, pp. 228-302.

[24] M. J. McTiffin, A. P. Hulbert, T. J. Ketseoglou, W. Heimsch, and G. Crisp, "Mobile Access to an ATM Network Using a CDMA Air Interface", IEEE J. Select. Areas Commun., vol. SAC-12, no. 5, June 1994, pp. 900-908.

[25] D. Raychaudhuri and N. D. Wilson, "ATM-Based Transport Architecture for Multiservices Wireless Personal Communication Networks", IEEE J. Select. Areas Commun., vol. SAC-12, no. 8, Oct. 1994, pp. 1401-1414.

[26] A. Baier, U-C. Fiebig, W. Granzow, W. Koch, P. Teder, and J. Thielecke, "Design Study for a CDMA-Based Third-Generation Mobile Radio System", IEEE J. Select. Areas Commun., vol. SAC-12, no. 4, May 1994, pp. 733-743.

[27] J. D. Bakker, R. Prasad, " Wireless Multimedia Communications Using a CDMA-Based TM-Air Interface", pp. 1128-1132.

[28] K. S. Gilhouson, I. M. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver, and C. E. Wheatly, " On the Capcity of a Cellular CDMA System", IEEE Trnas. Veh. Technol., vol. 40, no. 2, May 1991, pp. 303-312.

[29] J. W. Schwartz, J. M. Aein, and J. Kaiser, " Modulation Techniques for Multiple Access to a Hard-Limiting Satellite Repeater", Proc. IEEE, vol. 54, no. 5, May 1966, pp. 763-777.

[30] V. O. K. Li and X. Qiu, " Personal Communication Systems (PCS)", Proc. IEEE, vol. 83, no. 9, Sep. 1995, pp. 1210-1243.

[31] Y. Sanada, K. Seki, Q. Wang, S. Kato M. Nakagawa and V. K. Bhargava, "A Channel Equalization Technique on a Time Division Duplex CDMA/TDMA System for Wireless Multimedia Networks," IEICE Tran. Commun., vol. E78-B, no. 8, August 1995, pp. 1105-1115.

[32] Y. Sanada, K. Seki, Q. Wang, S. Kato M. Nakagawa and V. K. Bhargava, "A Transmission Power Control Technique on a Time Division Duplex CDMA/TDMA System for Wireless Multimedia Networks," IEICE Tran. Commun., vol. E78-B, no. 8, August 1995, pp. 1095-1102.

[33] J. Wu and R. Kohno, "A Wireless Multimedia Networks on a Time Division Duplex CDMA/TDMA," IEICE Tran. Commun., vol. E78-B, no. 7, July 1995, pp. 1002-1015.

[34] J. Blanz, A. Klien, M. NaBhan and A. Steil, " Performance of a Hybrid C/TDMA Mobile Radio System Applying Joint Detection and Coherent Antenna Diversity", IEEE J. Select. Areas Commun., vol. SAC-12 no. 4, May 1994, pp 568-579.

[35] N. Abramson, " Multiple Access in Wireless Digital Networks", Proc. IEEE, vol. 82, no. 9, Sep. 1994, pp. 1360-1369.

[36] N. Amitay, "Distributed Switching and Control with Fast Resource Assignment/Handoff for Personal Comunications Systems", IEEE J. Select. Areas Commun., vol. SAC-11, no. 6, pp. 842-849.

[37] A. S. Mahmoud, D. D. Falconer, and S. A. Mahmoud, "A Multiple Access scheme for Wireless Access to a Broadband ATM LAN Based on Polling and Sectored Antennas," IEEE J. Selec. Areas Commun., vol. SAC-14, no. 4, May 1994, pp. 596-608.

[38] E. Ayanoglu, K. Y. Eng, and M. J. Karol, "Wireless ATM: Limits, Challenges, and Proposals ", IEEE Personal Communications, August 1996, pp.18-34.

[39] R. L. Peterson, R. E. Ziemer, and D. E. Borth, Intoduction to Spread Spectrum Communications, New Jersey: Prentice-Hall, 1995.

[40] G. R. Cooper and C. D. McGillem, Modern Communications and Spread Spectrum, Singapore: McGraw-Hill, 1986.

[41] S. Verdu, "Minimum Probability of Error of Asynchronous Gaussian Multiple-Access Channel", *IEEE Trans. on Inform. Theory*, Vol. IT-32, No. 1, Jan. 1986, pp. 85-96.

[42] A. Duel-Hallen, J. Holtzman, and Z. Zvonar, "Multiuser Detection for CDMA Systems", *IEEE Personal Communications*, Vol.2, No. 2, April 1995, pp. 46-57.

[43] S. Moshavi, "Multi-user Detection for DS-CDMA Communications", *IEEE Communications Magazine*, Vol. 34, No. 10, Oct. 1996, pp. 124-136.

[44] S. Vedru, "Recent Progress in Multiuser Detection", in Advances in Communication and Signal Processing (Berlin: Springer Verlag, 1989).

[45] M.J. Juntti, and J.O. Lilleberg, "Implementation Aspects of Linear Multiuser Detectors in Asynchronous CDMA Systems", *Proc. of ISSSTA '96*, pp. 842-846.

[46] J. Lindner, "MC-CDMA and its Relation to General Multiuser / Multisubchannel Transmission Systems", *Proc. of ISSSTA '96*, pp. 115-121.

[47] J. G. Proakis, "Interference Suppression in Spread Spectrum Systems", *Proc. of ISSSTA '96*, pp. 259-266.

[48] A. Hottinen, and K. Pehkonen, "A Flexible Multirate CDMA Concept with Multiuser Detection", *Proc. of ISSSTA'96*, pp. 556-560.

[49] S. Vembu, and A. J. Viterbi, "Two Different Philosophies in CDMA - A Comparison", *Proc. of VTC' 96*, pp. 869-873.

[50] A. Hottinen, "Multiuser Detection for Multirate CDMA Communications", *Proc. of ICC' 96*, pp. 1819-1822.

[51] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, p. 385, McGraw Hill, 1991.

[52] R. G. Gallager, *Information Theory and Reliable Communication*, p. 337, John Wiley & Sons, Inc., New York, 1968.

[53] M. B. Pursley, "Performance Evaluation for Phase-Coded Spread-Spectrum Multiple Access Communication - Part I: System Analysis", *IEEE Tran. Commun.*, vol. Com-25, no. 8, Aug. 1977, pp. 795-799.

[54] K. Yao, Error Probability of Asynchronous Spread Spectrum Multiple Access Communication Systems", *IEEE Tran. Commun.*, vol. Com-25, no. 8, Aug. 1977, pp. 803-809.

[55] G. Turin, "The Effects of Multipath Fading on the Performance of Direct Sequence CDMA Systems", *IEEE J. Select. Areas Commun.*, vol. SAC-2, no. 4, July 1984, pp. 597-603.

[56] L. F. Chang, F. Ling, D. D. falconer, and N. Sollenberger, "Comparison of Two Orthogonal Coding Techniques for CDMA radio Communications Systems," *IEEE Tran. Commun.*, vol. Com-43, no. , 1995, pp.

[57] A. J. Viterbi, A. M. Viterbi, and E. Zehavi, "Performance of Controlled Wideband Terrestrial Digital Communication," *IEEE Tran. Commun.*, vol. Com-41, no. 4, April 1993, pp. 559-569.

[58] A. M. Viterbi and A. J. Viterbi, "Erlang Capacity of a Power Controlled CDMA System," *IEEE J. Select. Areas Commun.*, vol. SAC-11, no. 6, Aug. 1993, pp. 892-900.

[59] S. Ariyavisitakul and L.F. Chang, "Signal and Interference Statistics of a CDMA System with Feedback Power Control," *IEEE Tran. Commun.*, vol. Com-41, no. 11, Nov. 1993, pp.

[60] F. D. Priscoli and F. Sestini, "Effects of Imperfect Power Control and User Mobility on a CDMA Cellular Network," *IEEE J. Select. Areas Commun.*, vol. SAC-14, no. 9, Dec. 1996, pp. 1809-1817.

[61] J. M. Holtzman, "CDMA Power Control for Wireless Network," in Third Generation Wireless Information Network, S. Nanda and D. J. Goodman, Eds. Boston, MA: Kulwer, 1992, pp. 299-311.

[62] R. Lupas and S. Verdu, "Near-Far Resistance of Multiuser Detectors in Asynchronous Channels," *IEEE Tran. Commun.*, vol. Com-38, no. 4, April. 1990, pp. 496-508.

[63] M. Varanasi and B. Aazhang, " Multistage Detection in Asynchronous Code-Division Multiple-Access Communications," *IEEE Tran. Commun.*, vol. Com-38, no. 4, April. 1990, pp. 409-519.

[64] Boariu, A. and Ziemer, R. E., "Energy-Add Multistaage Detector: A Novel Multiuser Detection Method for Synchronous CDMA Systems," *IEEE J. Select. Areas Commun.*, vol. SAC-17, no. 11, Nov. 1999, pp 1953.

[65] A. Duel-Hallen, " A Family of Multiuser Decision-Feedback Detectors for Asynchronous Code-Division Multiple-Access Channels," *IEEE Tran. Commun.*, vol. Com-43, no. 2/3/4, Feb./Mar./April 1995, pp 421-434.

[66] E. S. Esteves and R. A. Scholtz, " Bit Error Probability of Linear Detectors in the Presence of Unknown Multiple Access Interference", Proc. of IEEE Globecom '97, pp. 599-603.

[67] H. G. Perros and K. M. Elsayed, " call Admission Control Schemes: A Review," *IEEE Communications Magazine*, Vol. 34, No. 11, pp. 82-91, November 1996.

[68] M. Schwartz, "Network Management and Control Issues in Multimedia Wireless Networks," *IEEE Pers. Commu.*, Vol. 2, No. , pp. 8-16, June 1995.

[69] E. W. Knightly and N. B. Shroff, " Admission Control for Statistical QoS: Theory and Practice," *IEEE Network Magazine*, pp. 20-29, March/April 1999.

[70] H. Saito, *Teletraffic Technologies in ATM Networks*, Artech House: Norwood, MA, 1994.

[71] T. Rapparpot, *Wireless Communications: Principles & Practice*, p. 385, Prentice Hall, 1996.

[72] A. Sampath and J. M. Holtzman, " Access Control of Data in Integrated Voice /Data CDMA Systems: Benefits and Tradeoffs," *Proc. IEEE ICC'97*, pp. 1271-1275, 1997.

[73] Z. Dzing, M. Jia and P. Mermelstein, "Adaptive Traffic Admission for Integrated services in CDMA Wireless Access Networks," *IEEE J. Select. Area Commun.*, Vol. 14, No. 9, pp. 1737-1747, Dec. 1996.

[74] J. Wu and R. Kohno, "A wireless Multimedia CDMA System Based on Transmission Power Control," *IEEE J. Select. Area Commun.*, Vol. 14, No. 4, pp. 1737-1747, May 1996.

[75] P. R. Larijani, J. W. Chinneck and R. H. Hafez, " Non-linear Power Control in Multimedia CDMA Wireless Networks," *IEEE Commun. Letters*, Sep. 1998.

[76] P. R. Larijani and R. H. Hafez, "A Frame work for Distributed access Control in Multimedia," to appear, *The Arabian Journal for Science and Technology.*

[77] M. Soroushnejad and E. Geraniotis, "Multi-Access Strategies for an Integrated Voice/ Data CDMA Packet Radio Network," *IEEE Tran. on Commun.*, Vol. 43, No. 2/3/4, pp. 934-945, Feb./March/April 1995.

[78] W-b Yang and E. Geraniotis, "Admission Policies for Integrated Voice and data Traffic in CDMA Packet Radio Networks," *IEEE J. Select. Area Commun.*, Vol. 12, No. 4, pp. 645-664, May 1994.

[79] T-K. Liu and J. A. Silvester, "Joint Admission/Congestion Control for Wireless CDMA Systems Supporting Integrated Services," *IEEE J. Select. Area Commun.*, Vol. 16, No. 6, pp. 845-857, August 1998.

[80] A. Sampath ,*et al*, "Power Control and Resource Management for a Multimedia CDMA Wireless System," *Proc of IEEE PIMRC' 95*, pp. 21-25.

[81] M. Soleimanipour, *et al*, "Modeling and Resource Allocation in Wireless Multimedia CDMA Systems," *Proc of IEEE VTC' 98*, pp. 1279-1283.

[82] C. Mihailescu, X. Lagrange and Ph. Godlewski, “ Radio Resource Management for packet Transmission in UMTS WCDMA System,” *Proc. of IEEE VTC' 99*, pp. 573-577.

[83] P. R. Larijani, R. H. Hafez and I. L. Lambadaris, “ Two Level Access Control Strategy for Multimedia CDMA,” *Proc of ICC' 98*, pp. 487-492.

[84] J. M. Jacobsmeyers, “Congestion Relief on Power-Controlled CDMA Networks,” *IEEE J. Select. Area Commun.*, Vol. 14, No. 9, pp. 1758-1761, December 1996.

[85] B. Hashem and E. Sousa, “Performance Evaluation of DS/CDMA Systems Employing Adaptive Transmission Rate under Imperfect Power Control,” *Proc of PIMRC '98*, Boston, USA.

[86] B. Hashem and E. Sousa, “A Combined Power/Rate Control Scheme for Data Transmission Over a DS/CDMA System,” *Proc of VTC '98*, Ottawa, Canada.

[87] R. D. Yates and C-Y Huang, “Integrated Power Control and Base Station Assignment,” *IEEE Trans. Veh. Technol.*, Vol. 44, No. 3, August 1995, pp. 638-644.

[88] A. Dabak and K. Kiasaleh, “ Integrated Voice and Video network Using Variable Rate CDMA,” *Proc. ICC '96*, pp. 1839-1843.

[89] J. capone and L. F. merakos, “Integrated Data Traffic into a CDMA Cellular Voice System,” *Wireless Networks*, Vol. 1, 1995, pp. 389-401.

[90] J. S. Evans and D. Everitt, “Effective bandwidth-Based Admission Control for Multiservice CDMA Cellular Networks,” *IEEE Trnas. Veh. Technol.*, vol. 48, no. 1, May 1999, pp. 36-46.

[91] Z. Lu, M. J. Karol, M. ElZarki, and K. Y. Eng, "Channel Access and Interference Issues in Multi-Code DS-CDMA Wireless (ATM) Networks," Wireless Networks, vol. 2, 1996, pp. 173-193.

[92] The cdma2000 ITU-R RTT Candidate Submission (0.18), Telecommunications Industry association (TIA), 1998.

[93] U. A. Baroudi and A. K. Elhakeem, "A Hybrid TDMA/MC-CDMA Utilizing Multiuser Detection for Integrated Wireless Networks," *IEICE Tran. on Commun.*, Vol. E83-B, No. 6, June 2000, pp.

[94] U. A. Baroudi and A. K. Elhakeem, " A New Hybrid CDMA/TDMA Multiuser Receiver System", *The 8th Virginia Tech/MPRG Symposium on Wireless Personal Communications*, Blacksberg, Virginia, June 10-12, 1998, pp. 111-121.

[95] U. A. Baroudi and A. K. Elhakeem, " A Novel Traffic Control in ATM-Based Wireless Networks", the IEEE Canadian Conference on Electrical and Computer Engineering, Waterloo, ON Canada, May 24-28, 1998, pp. 633-636.

[96] U. A. Baroudi and A. K. Elhakeem, "Adaptive Admission policy for Priority Queue Integrated CDMA Networks", the IEEE WCNC'99, New Orleans, USA, September 21-24, 1999, pp. 1266-1270.

[97] G. Anastasi et al., "MAC Protocol for Wideband Wireless Local Access: Evolution Toward Wireless ATM," *IEEE Personal Commun.*, pp. 53-64, October 1998.

[98] J. J. Bae, and T. Suda, "Survey of Traffic Control Schemes and Protocols in ATM Networks," *Proc. IEEE*, vol. 79, no. 2, Feb. 1991, pp. 170-189.

[99] A. J. Viterbi, *CDMA: Principle of Spread Spectrum Multiple Access Communication* (Reading, MA: Addison-Wesley, 1995).

[100] M. Schwartz, *Broadband Integrated Networks*, (Upper Saddle River, NJ: Prentice Hall PTR, 1996).

[101] S. M. Ross, *A Course in Simulation*, Second Ed., (San Diego, CA: Academic Press Inc., 1997).

Appendix A

Chernoff Bound on Error Probability under Rayleigh Fading multipath

A.1 Coherent Reception

We consider a DS-CDMA system with coherent pilot-aided demodulation with Rake receiver under Rayleigh multipath fading channel. Assume there are L paths components each with α_l amplitude and they could be estimated perfectly. Then the error probability becomes

$$\begin{aligned} P_E &= E\left[P_E(\alpha_1, \alpha_2, \cdots, \alpha_L)\right] < E\left[\prod_{l=1}^{L} \exp(-\alpha_l^2 N E_c / I_o)\right] \\ &= \prod_{l=1}^{L} E\left[\exp(-\alpha_l^2 E_s / I_o)\right] \triangleq \prod_{l=1}^{L} Z_l \triangleq Z \end{aligned} \tag{A.1}$$

Here, $E_s \triangleq N E_c$ is N-chip *symbol* energy. Let α_l be a Rayleigh-distributed random variable. Then the probability density function of α_l, is

$$p(\alpha) = \frac{2\alpha e^{-\alpha^2/\sigma_l^2}}{\sigma_l^2}, \qquad \alpha > 0 \tag{A.2}$$

Thus, for Rayleigh-distributed attenuation,

$$Z_l = E\left[e^{(-\alpha_l^2 E_s/I_o)}\right] = \int_0^\infty p(\alpha) e^{(-\alpha_l^2 E_s/I_o)} d\alpha \tag{A.3}$$

After some algebraic manipulation we get

$$Z_l = \frac{1}{1 + \sigma_l^2 E_s / I_o} \tag{A.4}$$

Finally, suppose that the L multipath components are *all Rayleigh of equal average strength*, so that $\sigma_l^2 = \sigma^2$, for all l

$$\overline{P_E} < Z \triangleq \prod_{l=1}^{L} Z_l = \left[\frac{1}{1 + \sigma^2 E_s / I_o}\right]^L = \left[\frac{1}{1 + (\overline{E_s}/I_o)}\right]^L \tag{A.5}$$

We may rewrite (A.5) as

$$\overline{P_E} < \exp(-\ln(1/Z)) \tag{A.6}$$

where

$$\ln(1/Z) = L \ln(1 + \overline{E_s}/I_o) \tag{A.7}$$

Now, for the case of unfaded Gaussian channel, it is easy to show that $\ln(1/Z) = E_s/I_o$. Therefore, from these two parameters, we can find the average excess energy (in decibels) required by the degraded channel to achieve the same performance as for unfaded signal in additive Gaussian noise. Figure A.1 shows this relation. For example, let E_s/I_o=4 dB that is required to achieve certain P_E under AWGN environment. Then, to achieve the same error performance under Rayleigh fading channel with $L = 2$, E_s/I_o=7 dB.

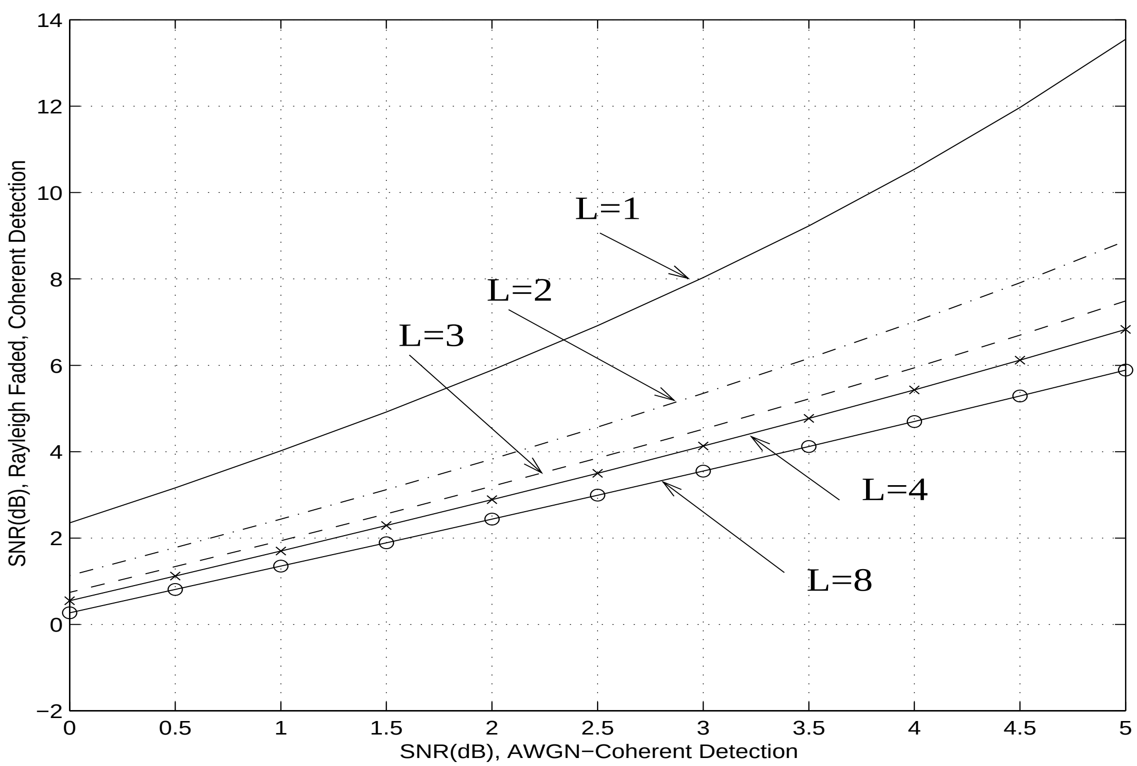


Fig. A.1: Required E_s/I_o for coherent demodulation of Rayleigh faded signals

A.2 Non-coherent Reception

The coherent reception considered above applies in the forward channel (one to many). However, in the reverse channel where the many-to-one scenario is applied, it is difficult to afford pilot signals that help in the process of coherent reception.

Therefore, non-coherent reception is considered for the uplink (reverse) channel.

In the results presented here, it is assumed that the timing for new paths is available, but phases and amplitude estimates are not. We can obtain a similar relation between the unfaded SNR and the excess average energy required for non-coherent faded signals to achieve the same error performance. Nevertheless, the analysis is much involved that the coherent case. Hence, we will only present the results by Fig. A.2.

Now, we can use these curves along with any other error performance curves under unfaded AWGN to find the excess energy required to achieve the same performance under Rayleigh fading channels. Assume that required P_E is 10^{-3}. From Fig. A.3, we find the required E_s is dB under unfaded Gaussian noise. Now, using Fig. A.1, we can map this value to find out how much power is needed to achieve the same P_E either for coherent reception or non-coherent reception. So find the following. For coherent reception, we need dB while for non-coherent reception we need dB.

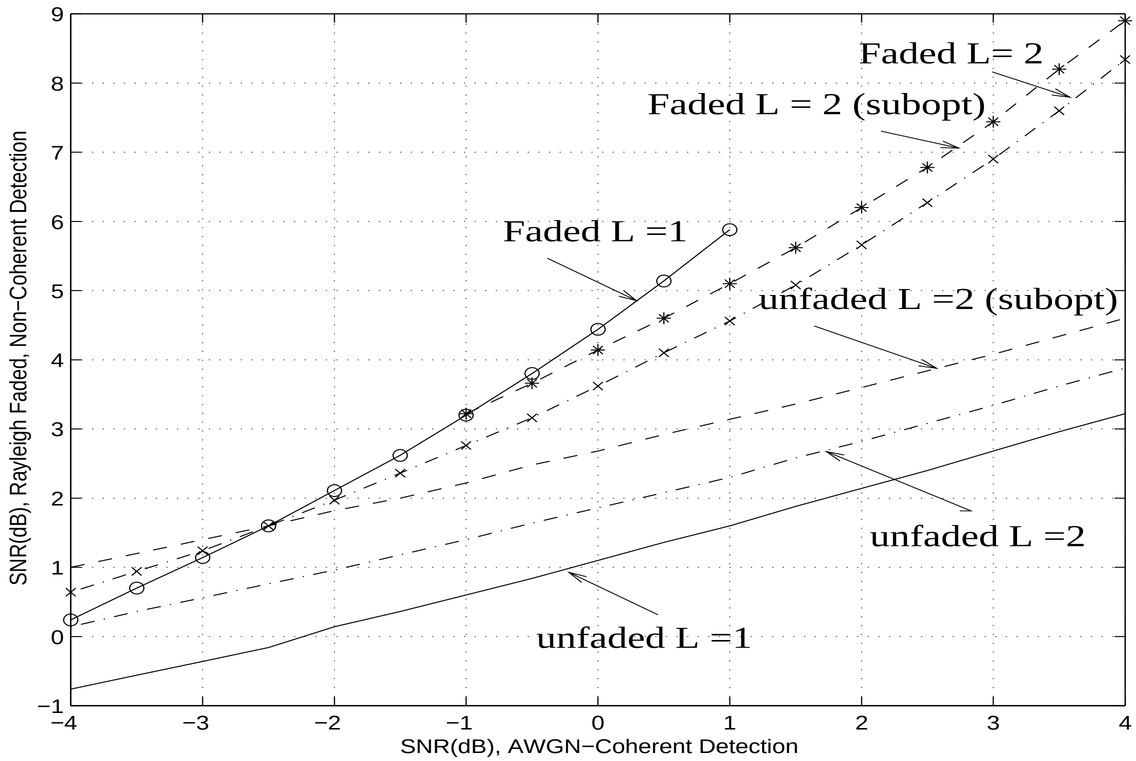


Fig. A.2: Required E_s/I_o for non-coherent demodulation of Rayleigh faded orthogonal signals (M=64)

Appendix B

Derivation of Non-Congested Buffer Probability

We want to find the probability that the incoming cell via uplink finds the buffer at the base station in a 'good' condition for at least q consecutive cell-time units. Let the probability that the buffer is not in a 'good' condition is $O_b(.)$, where (.) refers to the system parameters under which $O_b(.)$ has been evaluated. For convenience, denote the situation when the buffer is in a good condition by '1' (i.e. success) and otherwise by '0' (i.e fail). Hence, we can view the problem as a special case of W Bernoulli trials. So for a window of W cell-time units, we want to find all possible patterns of q or more consecutive 'success' events. In other words, we should find $S(.) = Pr(i \leq (W-q) \mid j \geq q)$ where i is the number of '0's (i.e. buffer not in a good condition) in the patterns and j is the number of consecutive '1's (i.e buffer in a good condition). Figure B.1 shows all possible patterns for $W = 5$ cells and $q = 3$.

$$S(.) \leq \sum_{j=q}^{W} \begin{pmatrix} W \\ j \end{pmatrix} (O_b)^{W-j}(1-O_b)^j \tag{B.1}$$

The global solution for all possible W and q values is very difficult because the patterns are not systematic. We have tried numerous different examples to figure our what formula that might be applicable for all possibilities and we could not reach for such formula. Nevertheless, a compact equation has been derived for certain range of W and q values.

To begin with, we are decomposing $S(.)$ into different terms where each term is composed of several elements that have common characteristics, i.e.

$$S(.) = \sum_{j=q}^{W} N_j (O_b)^{W-j}(1-O_b)^j \tag{B.2}$$

where N_j ($N_j \leq \begin{pmatrix} W \\ j \end{pmatrix}$) is the number of all possible patterns that have j ones and at least q of these ones are consecutive.

<u>Case1: Exactly q consecutive ones</u>

In this case, we have $W - q$ zeros. Imagine these q consecutive ones as one entity, then we can obtain $N_q = W - q + 1$ different possible patterns.

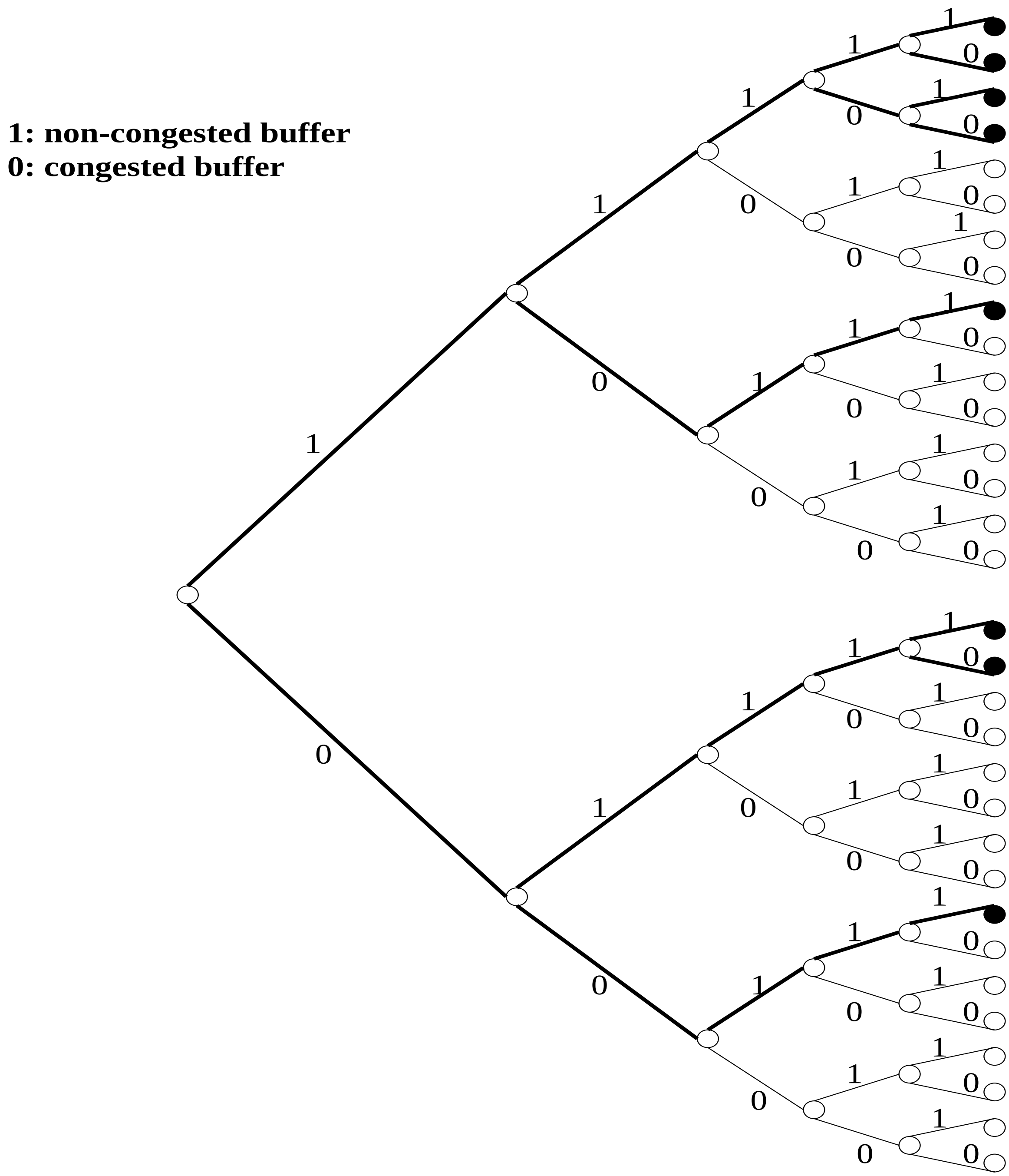


Figure B.1: Illustration of buffer congestion patterns over a measurement window; $W = 5$, $q = 3$

Case2: Exactly $q+1$ ones

Again consider the q consecutive ones as one entity, so we have $W-q$ spaces where the $(q+1)$th one can be in any position. Now, fix the position of the q ones, so there are

$$N_{q+1} = \begin{pmatrix} W \\ i \end{pmatrix} = W - q \tag{B.3}$$

distinct possibilities. Further, the group of q ones can take $W-q$ different positions. Therefore, the number of distinct possible patterns is $(W-q)^2$.

Now, The question is can we continue with the same procedure to find these probabilities? The answer is NO.The reason is that there is no systematic approach for all possible patterns. However, the following proposition may lead to more precise value of $S(.)$ compared to Eq. B.1.

Proposition

For any pattern composed of N ones and M zeros, where $W = M + N$, the necessary condition for guaranteeing all permutations (i.e. $\begin{pmatrix} W \\ N \end{pmatrix}$) have at least q $(q \leq N)$ consecutive ones is

$$\frac{N}{M} \geq q \tag{B.4}$$

Therefore, we can write $S(.)$ as follows

$$\begin{aligned} S(.) &= (W-q+1)(O_b)^{W-q}(1-O_b)^q \\ &+(W-q)^2(O_b)^{W-q-1}(1-O_b)^{q+1} \\ &+\cdots \\ &+\sum_{i=K}^{W} \begin{pmatrix} W \\ i \end{pmatrix} (O_b)^{W-i}(1-O_b^1)^i \end{aligned} \tag{B.5}$$

From the above proposition, the last term to be hold the following inequality should hold, i.e.

$$\frac{K}{W-K} \geq q \tag{B.6}$$

It is obvious from B.6 that some terms are not defined because the condition in B.4 does not hold. Therefore, to simplify the analysis, we chose to assume these missing terms as if they exist and then we end with a lower bound probability of $S(.)$, i.e.

$$\begin{aligned} S(.) \leq\ & (W-q+1)(O_b)^{W-q}(1-O_b)^q \\ & +(W-q)^2(O_b)^{W-q-1}(1-O_b)^{q+1} \\ & +\sum_{i=q+2}^{W}\begin{pmatrix} W \\ i \end{pmatrix}(O_b)^{W-i}(1-O_b^1)^i \end{aligned} \tag{B.7}$$